Oikos 000: 001–011, 2015
doi: 10.1111/oik.02305

Subject Editor: Francois Massol. Editor-in-Chief: Dries Bonte. Accepted 26 June 2015

Exploring the evolutionary signature of food webs' backbones using functional traits

Giulio V. Dalla Riva and Daniel B. Stouffer

G. V. Dalla Riva (orcid.org/0000-0002-3454-0633)(gvd16@uclive.ac.nz), School of Mathematics and Statistics, and School of Biological Sciences, Univ. of Canterbury, Room 523 - Erskine Building Science Road Ilam, Christchurch, Canterbury 8041, New Zealand. – D. B. Stouffer, Centre for Integrative Ecology - School of Biological Sciences Univ. of Canterbury, Christchurch, Canterbury 8041, New Zealand.

Increasing evidence suggests that an appropriate model for food webs, the network of feeding links in a community of species, should take into account the inherent variability of ecological interactions. Harnessing this variability, we will show that it is useful to interpret empirically observed food webs as realisations of a family of stochastic processes, namely random dot-product graph models. These models provide an ideal extension of food-web models beyond the limitations of current deterministic or partially probabilistic models. As an additional benefit, our RDPG framework enables us to identify the pairwise distance structure given by species' functional food-web traits: this allows for the natural emergence of ecologically meaningful species groups. Lastly, our results suggest the notion that the evolutionary signature in food webs is already detectable in their stochastic backbones, while the contribution of their fine wiring is arguable.

The existence of an interaction between two species in a food web has often been regarded as a static and largely deterministic event (May 2006): if an individual predator from species *i* has been observed to consume an individual prey from species *j*, this link is thought to occur everywhere species *i* and *j* co-occur. A growing body of evidence, however, challenges this view and supports the notion that food webs are inherently dynamic and the product of many events, some of which exhibit a stochastic behaviour (Holling 1973, Coulson et al. 2004, Mullon et al. 2009, Black and McKane 2012). For example, the probability of observing an interaction between two species in a certain location depends on a mixture of neutral and niche processes, in addition to other behavioral and environmental factors (Fortuna et al. 2013, Canard et al. 2014, Poisot et al. 2014). More specifically, interactions depend at least on species' local abundances – determining the encounter probability – and on species' local phenotypes as characterized by their trait values – determining the interaction intensity (Poisot et al. 2014).

Beyond specific trait values, a species' ensemble of predators and prey can be regarded as an emergent form of its local phenotype – its 'trophic niche'. Within the food-web literature, it is widely acknowledged that the importance of the various traits in determining the species trophic niche is not uniform (Petchey et al. 2008a). The traits playing a major role in shaping the trophic niche are known as 'trophic traits'; in particular, body size is commonly assumed to be the most important (Jennings et al. 2002). It has also been shown that many emergent properties of food-web structures can often be effectively predicted by models based on just two traits: a predator's body size and the range of body size of that predator's prey (Williams and Martinez 2000, 2008, Stouffer et al. 2005, 2011, Williams et al. 2010, Zook et al. 2011, Gravel et al. 2013), though this may depend on ecosystem type (Woodward et al. 2005a, Stouffer et al. 2011). However, the performance of many standard food-web models is reduced when we consider their ability to predict single interactions (Petchey et al. 2008a). This decrease in performance likely arises because of the models' deterministic nature (Williams et al. 2010, Gravel et al. 2013) and the phenomenological way in which they relate to species' traits (Stouffer 2010, Eklöf et al. 2013).

We detail an approach here to the study of food webs as intrinsically stochastic processes, by modelling food webs as directed random dot-product graphs (RDPG). If we shift from the deterministic focus, one can consider every observed food web as the outcome of a chain of three distinct processes: first, ecological and evolutionary factors determine the morphology of species, which is the set of traits of each species; second, the species' traits – or, more precisely, a species' specific subset of those traits – determine which interactions can occur; third, an observer detects some, or all, of the species' interactions. Each of these steps involves elements of stochasticity. In our modelling framework here, we focus in particular on the second step – linking traits to interactions – and set aside the issues related to the other two processes. We do not assume a priori knowledge of the number and identity of species' traits that determine species' trophic niches or interactions. Instead, we estimate species' functional traits through a singular value decomposition of

the observed food web's adjacency matrix. In doing so, our framework is able to shed insight into the long-standing question of the dimension of niche space in ecological communities (Cohen 1977, 1978, 1983, Morowitz 1980, Cohen and Palka 1990, Stouffer et al. 2006, Eklöf et al. 2013). Our approach also provides a novel perspective of species' trophic similarity in terms of their relative functional roles (Allesina and Pascual 2009, Jordán 2009, Stouffer et al. 2012, Eklöf et al. 2013). Lastly, we find that food webs can be efficiently described via low dimensional traits and that they exhibit an evolutionary imprint which is due mostly to the structure of food webs' stochastic backbone.

Methods and material

Empirical food webs

We applied our methodological approach to nine different food webs, widely varying in location, composition and species' community size. For the majority of our analyses and results, we will focus on the two largest food webs. The first large web was compiled for the Serengeti National Park (Baskerville et al. 2011), and it is made up of 161 species and 592 feeding relationships. Amongst those 161 species, 129 are plants, 23 are herbivores and 9 are carnivores. Most of the links (507) are between herbivores and plants whereas 85 are between animal species. The second large web is a highly resolved food web for the antarctic Weddell Sea (Jacob et al. 2011). This food web is composed of 488 taxonomically identified species, four distinct non-living source nodes (e.g. detritus), and features more than 16 000 predator–prey interactions. We chose these two food webs because they are both well resolved to the species level which then allows for a robust phylogenetic analysis. We expect that the differences in their latitude, their environment, and their species composition would help ascertain the utility of our approach. The remaining seven food webs are smaller and were compiled by different authors (Closs and Lake 1994, Dawah et al. 1995, Memmott et al. 2000, Woodward and Hildrew 2001, Harper-Smith et al. 2005, Jonsson et al. 2005, Woodward et al. 2005b, Ledger et al. - in Petchey et al. 2008b). We analysed the latter in order to propose a more complete comparison of the model we are introducing with other well recognized models (Petchey et al. 2008b, Allesina and Pascual 2009, Rohr et al. 2010). The sizes of these seven webs vary from 25 to 80 species.

Random dot product graphs

Sociologists are often faced with the problem of predicting the presence and absence of unobserved interactions in a community of individuals where only some of the interactions have been observed. A classic approach to the problem is based on the characterization of each individual in terms of its features – e.g. interests, hobbies, acquaintances. Then, the probability of establishing a relationship between two individuals is let to be given by the similarity of their features.

In this scenario, random dot product graph (RDPG) models are an effective solution to the task of inferring individuals' features given their observed or established interactions (Wasserman 1994, Nickel 2007, Young and Scheinerman 2007). In such an RDPG model, each individuals' features are mathematically expressed using a vector of traits, and the interaction probability between individuals is a function of the dot product of their trait vectors. As such, the probability of an interaction increases as the two vectors approach each other – the angle between them decreases – and is largest when they are collinear – the angle between them is zero.

Directed random dot product graphs

As many social relationships are symmetric, the underlying interaction graph generated by a sociological RDPG model is undirected. On the other hand, the ecological relationships we are interested in (e.g. predation) are usually not symmetric. Therefore, we must consider a directed RDPG (Young and Scheinerman 2008) to model food webs. This necessitates that species be described not by a single vector of traits but by a pair of vectors, which we will refer to as their 'foraging functional trait' and the 'vulnerability functional trait' vectors. The foraging functional traits of species i will help determine the probability of observing links toward i – the behavior of i as a predator (or consumer) — whereas the vulnerability functional traits of species i will help determine the probability of observing links from i — the behavior of i as a prey (or resource). As with undirected RDPGs, the probability of a link from species j to species i (i.e. of i consuming j) will be given by the dot product between the vulnerability functional traits of j and the foraging functional traits of i.

To better understand the theory behind directed RDPGs, consider three hypothetical species a, b and c as in Fig. 1. Here, each species is associated with a pair of two-dimensional functional traits (in this case, the x- and y-coordinates). For the sake of simplicity in this example, we have imposed their functional trait vectors to have length equal to 1 so that the difference between them is given just by their angular distances. Moreover, we have constrained all of them to be situated in the positive quadrant so that the cosine of the angles between falls in the interval [0, 1].

In this example, the angle between the vectors $a^{(f)}$ and $b^{(v)}$ is smaller than the angle between the vectors $a^{(f)}$ and $c^{(v)}$; mathematically, this implies that the dot product of the vectors $a^{(f)}$ and $b^{(v)}$ is greater than the dot product of the vectors $a^{(f)}$ and $c^{(v)}$. Within the RDPG framework, this also implies that there is a greater probability of observing a link from b to a – a consuming b – than a link from c to a. Similarly, we can see that the angle between $c^{(f)}$ and $a^{(v)}$ is larger than the angle between $c^{(v)}$ and $a^{(f)}$.This again implies that there is a greater probability of observing a link from c to a than a link from a to c. Considering the pairwise angular distances of all other species, we can directly infer the most likely structure of the three-species food web.

Estimating species' functional traits

The ensemble of all species' traits vectors determine the probability of observing each and every potential link in the food web. As a result, they determine the probability of

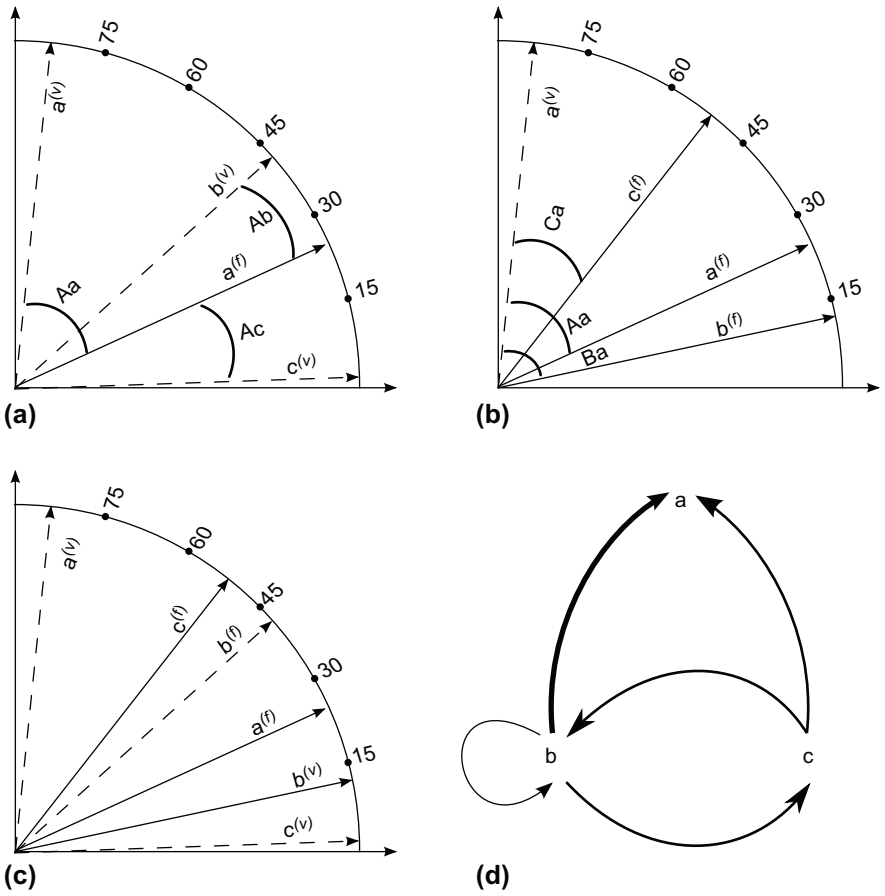

Figure 1. The functional trait space of three hypothetical species and their corresponding food web. (a) The foraging functional traits of species *a* and the angular distance with the vulnerability functional traits of species *b* and *c*. (b) The vulnerability functional traits of species *a* and the angular distance with the foraging functional traits of species *b* and *c*. (c) The foraging and vulnerability functional trait vectors for all the species. (d) The most likely food web, where the width of each interactions is proportional to the probability of the interaction, i.e. the angle between the corresponding foraging and vulnerability functional trait vectors.

sampling a certain food web from the space of all allowable realizations (i.e. its likelihood). Of course, in practice, we can never directly observe the process that generates a food web; instead, we usually obtain a single sample from the set of all the possible food web realizations. This implies that we then need to estimate species' foraging and vulnerability functional traits that are most likely to have produced the observed food webs.

For each given dimension d of the RDPG model, the most likely functional traits are those that minimize the distance between the observed adjacency matrix and the matrix whose entries are the probabilities $p(i \rightarrow j)$ given by the RDPG model (with those functional traits). We perform this estimation as follows. Let G be an observed food web composed of S species and A_G its $S \times S$ adjacency matrix. Finding the d-dimensional traits that minimize the distance between the model matrix and A_G is equivalent to finding a pair of $S \times d$ matrices and such that they minimize the distance between A_G and the product $\hat{L} \rightarrow \hat{R}^t$. Random dot product graph theory (Lyzinski et al. 2013, Tang et al. 2013) provides an efficient algorithm to solve this problem based on singular value decomposition (SVD). Here, there is a strong parallel with principle component analysis, in the sense that a high-dimensional dataset (in our case the adjacency matrix) is reduced to a lower-dimensional dataset (the functional traits), with a minimal loss in the information content.

To do this, we first obtain an SVD of A_G into three matrices L, $\sum$, R, such that $A_G = L \times \sum \times R^t$. Here, L and R are real, orthogonal $S \times S$ matrices, and $\sum$ is an $S \times S$ diagonal matrix whose non-decreasing ordered entries are the singular values of A_G. With the SVD of A_G, we can then define three new matrices that capture the d leading traits: 1) L', an $S \times d$ matrix given by the first d columns of L; 2) R', an $S \times d$ matrix given by the first d columns of R; and 3) $(\sum)^{1/2}$, an $d \times d$ diagonal matrix defined by the square root of the d greatest singular values of A_G. From these reduced matrices, $L\hat{}$ is given by $L\hat{} = L' \times (\sum)^{1/2}$ and $R\hat{}$ is given by $R\hat{} = (\sum)^{1/2} \times R'$. The rows of $L\hat{}$ and $R\hat{}$ give the species' vulnerability functional and foraging functional traits, respectively. Note that these traits are not uniquely identifiable, as any transformation of the matrix $L\hat{}$ and $R\hat{}$ preserving their dot product would be acceptable.

Choosing the trait dimensionality

The dimension of an RDPG model has direct effects on the variability of the food webs the model produces; that is, a higher dimension corresponds to a lower variance of the

estimated food webs. To explain this notion further, let L and R be specified, S^2-dimensional matrices such that $L \times R^t = A$ is a binary matrix (one can see them as the full rank traits estimated from A). For each $d < S$, let us defined a RDPG model with functional traits given by L_d and R_d, the first d columns of L and R. As d increases, the variance of the probability distribution given by the RDPG model decreases and the sampled food webs will share more and more links with A. Hence, a pivotal element of the RDPG approach to food webs is the identification of the model dimensionality.

Ideally, one would infer the variability of a food web from empirical data. This, however, is not always possible, and repeated observations of food webs are rare. An alternative is given by a graph-theoretical approach: the decreasing sequence of singular values of the observed food web's adjacency matrix A_G is of great utility in assessing the presence of structure in the graph, and hence in delimiting an appropriate dimensionality interval (Chatterjee et al. 2014).

Along these lines, different methods to asses a suitable model dimensionality have been proposed in the statistical literature. A first, conservative upperbound for d is given by the number of non-zero singular values, $\sum_+$. This is because every coordinate i after $\sum_+$ is strictly null. Moreover, each additional coordinate contributes proportionally to the i-th singular value gap, that is the distance between the i-th and the $(i+1)$-th singular values (Andrews and Patterson III 1976).

To do so, we investigate a variety of methods. We perform an exploratory data analysis by looking at the scree plot of the singular values. The objective then is to identify an 'elbow' in the data: we expect the sequence to decrease quickly up to a certain value of d, after which the decrease will be noticeably slower (Cattell 1966). Cattell's approach, although widely used, has the drawback of depending on a personal judgment. Note, however, that we are not trying to identify the optimal d but rather an interval of acceptable values for it.

We complemented our ocular intuitions with two 'researcher independent' methods that try to estimate an optimal d. The first method maximizes a profile likelihood function, and was developed by Zhu and Ghodsi (2006) in the scenario of PCA analysis. The second method is based on the identification of a universal singular value threshold: that is, the identification of a threshold such that, considering only those coordinates associated with singular values higher than the threshold, the distance between the estimated matrix and the 'real' matrix (the matrix given by the 'real' model) is asymptotically small (Chatterjee et al. 2014, Gavish and Donoho 2014). This latter approach incorporates the structural hypothesis used in the model (i.e. different random-graph models have different threshold values).

Assessing model performance

To further corroborate our choice of dimensionality, we observe the model performance as a function of the trait-vector length d. Specifically, we assessed the RDPG model's fitting performance in two distinct ways: 1) through its sensitivity, i.e. the ratio of correctly predicted links to observed links, and 2) through its accuracy, i.e. the ratio of correctly predicted observed links and correctly predicted absent links to the squared size of the community. To compute these ratios for each model dimensionality d within our estimated optimal interval, we sampled food webs such that each link had an independent probability of being observed (given by the d-dimensional model). Notice that, for the reasons illustrated in the previous section, we expect the performance to grow with d.

Next, we assessed the RDPG model's predicitve performance in a leave-one-out cross validation test, and we again focused on the sensitivity (correctly predicted links, the 'ones' in the observed adjacency matrix), specificity (correctly predicted absent links, the 'zeros' in the observed adjacency matrix) and accuracy (correctly predicted entries in the adjacency matrix). In the leave-on-out procedure, we sequentially treated each element of the adjacency as unobserved (i.e. absence of observation, not observation of absence). To do so, we set the entry equal to the a priori probability p of observing an interaction in the food web – that is, the connectivity of the food web without that interaction. Then, we estimated the d-dimensional functional traits on the modified adjacency matrix and computed the a posteriori probability p_{ij} of observing an interaction corresponding to that entry. We classified an entry with value greater than 0.5 as presen and an entry with value less than 0.5 as absent. Notice that this is equivalent to averaging the presence/absence of a link over a large sample of randomly sampled food webs where each links is observed with a probability equal to p_{ij}. We compared the model-estimated food web, which depends on d, with the originally observed one.

We computed the Akaike information criterion value for the RDPG model on all the food webs we analysed. Specifically, the AIC of our model is given by:

$$2\cdot(S\cdot 2\cdot d) - 2\cdot\left(\sum_{ij|A_{ij}=0}\log\left(1-p_{ij}\right)+\sum_{ij|A_{ij}=1}\log\left(p_{ij}\right)\right) \qquad (1)$$

where p_{ij} is the probability specified by the model for an interaction from species i to species j, A_{ij} is the entry corresponding to the interaction from species i to species j in the adjacency matrix, S is the number of species, and d is the length of the trait vectors. In order to compare the RDPG to the other models, we obtained the fitting performances for the other models – for all but the two largest webs – from the literature (Allesina and Pascual 2009, Rohr et al. 2010, Allesina 2011). For the Serengeti food web, we obtained the species allocation in the 14 groups from the literature (Baskerville et al. 2011) and then estimated link density between each pair of groups ourselves. We computed the likelihood of the group model as described above for the RDPG model, considering the number of parameters equal to $S + \gamma^2$; that is, the number of species plus the square of the number of groups (but see the caveat in Allesina 2011). All further details regarding the implementation and fitting performance of these models can be found in the original publications Petchey et al. (2008a), Allesina and Pascual (2009), Rohr et al. (2010) and Baskerville et al. (2011).

Phylogenetic signal

For the two largest webs, we lastly explored the phylogenetic signal of the observed food webs' RDPG approximations as

a function of the model dimensionality d. Though, strictly speaking, the upperbound of our analysis is arbitrary, note that it is much less than the full rank of the food webs' adjacency matrices.

We quantified the presence of a phylogentic signal by comparing the phylogenetic variance–covariance matrix (Revell et al. 2008) between species in a community with the dissimilarity matrix obtained by considering the pairwise Jaccard similarity (Real and Vargas 1996) computed from that community's adjacency matrix (Rohr and Bascompte 2014) (as sampled from a d dimensional model or testing single dimensions). The Jaccard similarity of two species in a food web is the number of common predators that consume both focal species divided by the number of predators that consume at least one of the two. Across all pairs of species, this defines a pairwise similarity matrix depending on the model dimension. Similarly, one can also compute the Jaccard similarity based on common prey, or on common predators and prey. For each model dimension considered, we computed the correlation between 99 sampled Jaccard similarity matrices (for species as prey, predators, or both) and the phylogenetic variance–covariance matrix, and we tested for significance using a Mantel-test with 999 randomizations.

For the Serengeti National Park data, we used a dated phylogenetic tree based on molecular data compiled by De Zwan (Supplementary material Appendix 1). For the Weddell Sea, we approximated the real phylogeny via a cladogram obtained from the taxonomic classification of the species as given by the Integrated Taxonomic Information System (<www.itis.gov>; information retrieved on 2014-11-11). Given these trees, we estimated the phylogenetic variance–covariance matrix under the assumption of Brownian motion trait evolution (Felsenstein 1985, Pagel 1992). The model assumes that the traits evolved as independent identically distributed Brownian motion along the lineages defined by the phylogenetic tree.

Results

Trait dimensionality

Our exploratory analysis, identified an optimal model dimension upperbound of $d = 6$ for the Serengeti food web and of $d = 8$ for the Weddell Sea food web (Fig. 2). The Zhu and Ghodsi method indicates a smaller upperbound of $d = 3$ for the Serengeti food weband of $d = 6$ for the Weddell Sea food web. Alternatively, the universal singular value threshold method indicates an upperbound of $d = 4$ (or $d = 2$ for the hard singular value threshold) for the Weddell food web, while it failed to indicate a upperbound for the Serengeti food web (the threshold is higher than all singular values).

The upper bound found with the singular value threshold methods for the smaller webs is consistently $d = 1$ or less (the threshold is higher than all singular values). For these webs, the upper bound found with Zhu and Ghodsi's method is consistently $d = 3$ or lower, except for the Grassland food web (Dawah et al. 1995) for which $d = 8$. All the sequences of singular values and the upperbounds identified by the different methods are presented in the Supplementary material.

Model performance

Fitting performance

The fitting performance in terms of sensitivity was high: more than 60% in a three dimensional model, in both the Serengeti's and Weddell's food webs, and more than 80% in a six (eight) dimensional model in the Serengeti (Weddell)

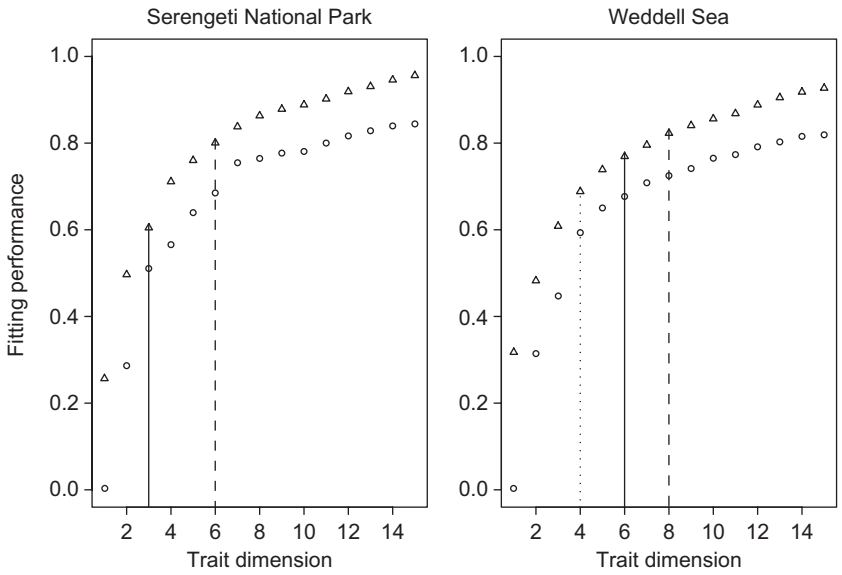

Figure 2. Model fitting performance varies with model dimensionality. We show the cumulative sum of the singular value gaps of the food web's adjacency matrix (dots) and fitting sensitivity (triangles) as a function of functional traits' dimension for the Serengeti National Park and Weddell Sea food webs (left and right, respectively). The dotted line corresponds to the dimensionality suggested by the Universal Singular Threshold method, the solid line to Zhu and Ghodsi's method, and the dashed line to our visual examination. Comparable figures are offered in the Supplementary material for the smaller webs.

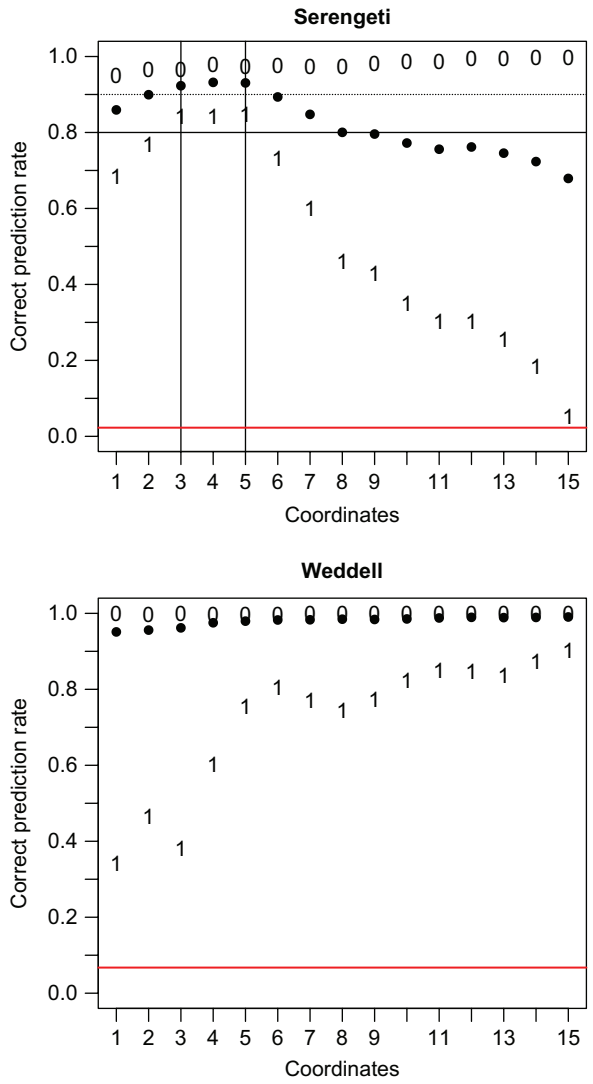

Figure 3. Model predictive performance varies with model dimensionality. We show the predictive performance as a function of functional traits' dimension for the Serengeti National Park and Weddell Sea food webs (top and bottom, respectively), for observed links, (1) non-observed links (0), and all pairs of species (full dots). The bottom red line indicates the predictive power of a null model (each link has independent probability of being observed equal to the food web connectance). In the Serengeti plot, we have highlighted the region of peak predictive performance between the two vertical lines. Comparable figures are offered in the Supplementary material for the smaller webs.

food web. The fitting performance in terms of accuracy was even higher: more than 95% with $d = 3$ in both food webs. On the smaller food webs, the accuracy was consistently above 80% while the sensitivity was more variable, ranging between 20% and 80% at $d = 1$. It was, however, consistently above 80% starting at dimensionalities between $d = 3$ and $d = 8$, with the exception of the Tuesday Lake food web (for which the sensitivity never reached 80%).

Predictive power

The predictive power of the RDPG model based on the leave-one-out analysis was high, see Fig. 3. In the Serengeti's food web more than 60% of the observed links were correctly predicted for models with dimension in the range $d \in \{1,...,7\}$ and more than 80% of the observed links were correctly predicted for $d \in \{3, 4, 5\}$.The performance in terms of accuracy was even higher: more than 95% with $d = 3$ in both food webs. Both accuracy and sensitivity were high on the Weddell sea food web. We could identify a saturating trend in this dataset as well, although we could not detect a peak for values of d between 1 and 16 (and we could not extend our analysis further for computational reasons). Nevertheless, we expect a similar overall trend to be present in this case as well.

Performance comparison

The RDPG performed well compared to the other three models we analysed in terms of both fitted and predicted linkwise accuracy. With $d = 1$, the RDPG model's accuracy already exceeded the accuracy of Rohr et al.'s (2010) and Petchey et al.'s (2008a) models for six of the seven smaller webs; in the case of the Broadstone stream food web, this is true starting from $d = 2$ (with $d = 1$ the RDPG model's accuracy roughly matched that of Rohr et al.'s model). The linkwise sensitivity and accuracy of the blockmodel proposed by Allesina and Pascual (2009) is outperformed by the RDPG model on the Serengeti food web starting from $d = 3$ and $d = 2$, respectively. A graphical summary of these comparisons is offered in the Supplementary material.

When amenable to comparison, the RDPG model had a lower AIC than Petchey's allometric diet breadth model and an AIC that was marginally lower or higher than that of Rohr's model and Allesina and Pascual's model (Table 1). As one can see from Table 2, this is most likely due to the number of parameters, which is considerably higher in the RDPG model than in any other model. For Broom (Memmott et al. 2000) and Grassland (Dawah et al. 1995), two of the smallest webs, the RDPG model assigns a null probability to (at least one) of the observed interactions when d was low. Hence, its log-likelihood in these situations is minus infinity and its AIC is plus infinity. See the Supplementary material for more details.

Species' functional traits

The distribution of species in the functional-trait space can help us explore their ecological role. It would be particularly

Table 1. AIC scores for a directed random graph (Erdös and Rényi 1960), Petchey's allometric diet breadth (as reported in Allesina 2011), Rohr's (Rohr et al. 2010), and Allesina and Pascual's (Allesina, 2011, but see Appendix C therein for a caveat about using AIC in this model), and the RDPG model (for the trait length d that minimises the AIC score). The values highlighted by * are computed here on the basis of the data published in (Baskerville et al. 2011, however the caveat discussed in Allesina, 2011 holds here as well).

	Random	Petchey's	Rohr's	Group	RDPG	d
Broadstone Stream	809	811	285	272	298	1
Broom	974	1111	657	653	Inf	–
Coachella Valley	866	777	–	411	445	3
Grasslands	1007	–	944	–	Inf	–
Mill Stream	2813	2641	1358	1222	1275	2
Skipwith Pond	2529	2654	1491	1360	1485	2
Tuesday Lakes	2893	2513	873	833	1418	3
Serengeti	5647*	–	–	2478*	3416	3

Table 2. Log-likelihood of a random graph (Erdös and Rényi 1960), Allesina's and Pascual's (Allesina and Pascual, 2009), and the RDPG model (for the trait length d that minimises the AIC). *from our computation. We omitted the two food webs where we could not choose the RDPG dimension based on AIC. (the log-likelihoods for all d in {1, . . . , 25} can be found in the Supplementary material).

	Random	Group	RDPG
Broadstone Stream	−403.362	−70.941	−91.12
Coachella Valley	−432.108	−115.637	−66.69
Mill Stream	−1405.41	−466.776	−477.4
Skipwith Pond	−1263.357	−488.002	−458.2
Tuesday Lakes	−1445.359	−199.537	−271.0
Serengeti	−2822.5*	−881.3745*	−742.2*

useful in detecting 'outliers', i.e. species with a truly unique role in food web (Petchey et al. 2008b, Jordán 2009), and clumps, i.e. species with a similar food-web role (Allesina and Pascual 2009, Stouffer et al. 2012). Along these lines, we found that the strongest differentiation in the Serengeti food web was between animals and plants. A phenomenon clearly visible in the first coordinate of foraging functional and vulnerability functional traits (rightmost column of Fig. 4). Top predators and grazers were spread far from the origins of the coordinate axis while plants are stacked upon the axis' origin. We would fully expect this behavior as they do not have any incoming links; that is, they do not feed on any other species (included in this food web).

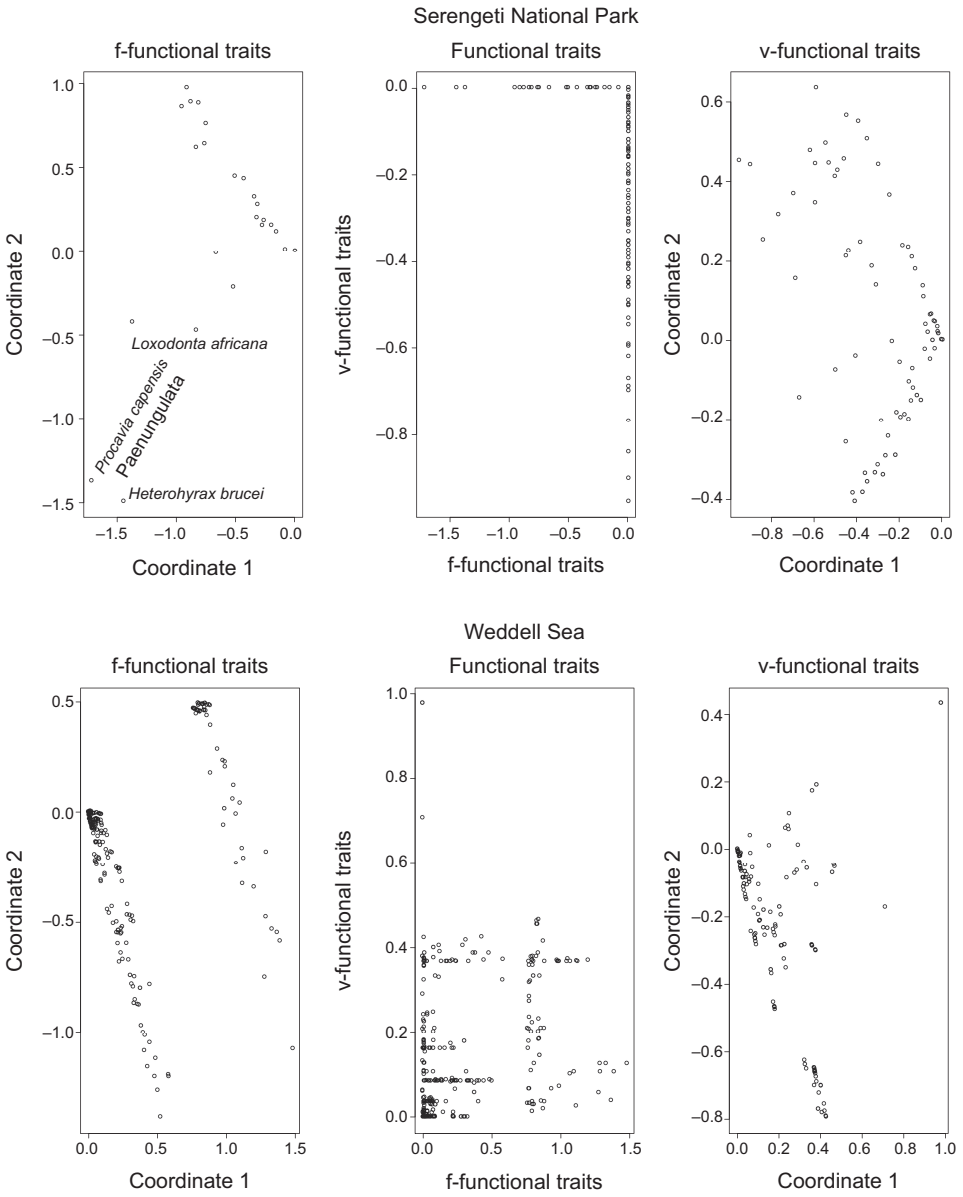

Figure 4. The distribution of functional traits for species in the Serengeti National Park and Weddell Sea food webs. In the left column we show the first two coordinates of the foraging functional traits, and in the right column we show the first two coordinates of the vulnerability functional traits. In the middle, we show the first foraging functional-trait coordinate against the first vulnerability functional-trait coordinate. We can notice the outlier position of the Hyracoidea (and of *Loxodonta africana*, their closest evolutionary relative) in the Serengeti National Park. The deep distinction between plants and animals in the Serengeti is also visually apparent (central panel).

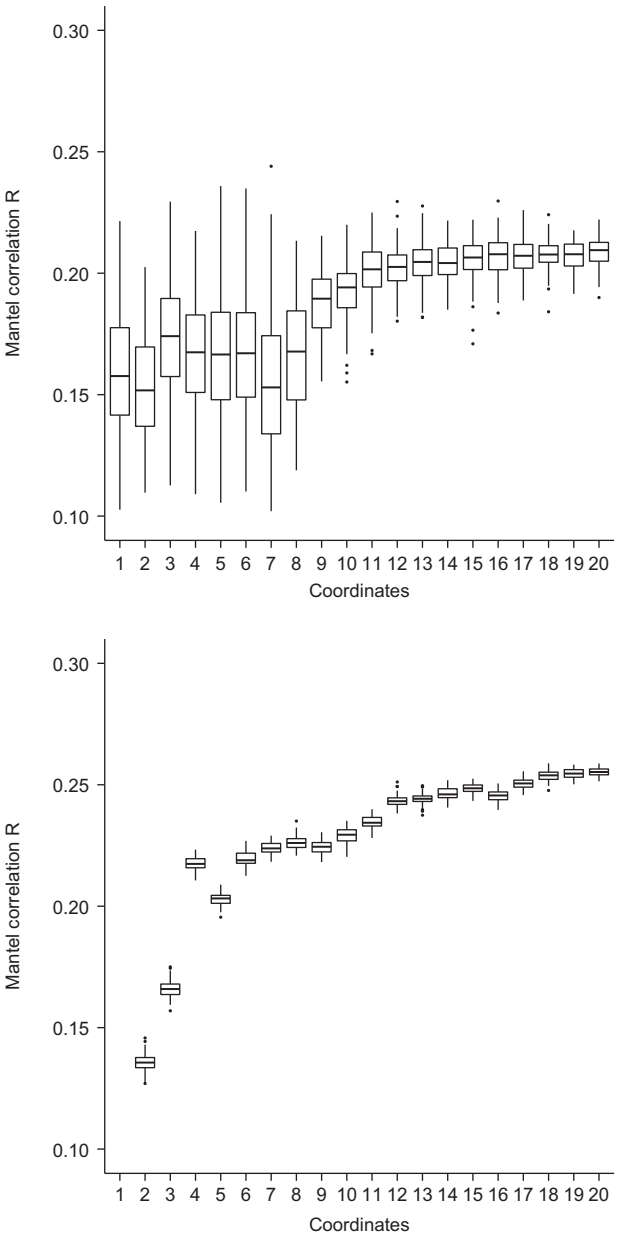

Figure 5. Correlation between the similarity of species' roles (as both predator and prey) and the phylogenetic variance–covariance matrix for the Serengeti National Park (top) and Weddell Sea (bottom) food webs. In both panels, we show the (significant at the level $p < 0.01$) correlation for 99 sampled food webs as a function of the length of species' trait vectors. Notice that the model with $d = 1$ for the Weddell Sea is not significant and we don't show the correlation value.

Similarly, two Hyracoidea species (*Heterohyrax brucensis* and *Procavia capensis*) appeared unique in the foraging functional trait space of the Serengeti web, suggesting that their behaviour as predators is 'peculiar'. This unique role of Hyracoidea was also observed by the species grouping proposed by Baskerville et al. (2011) based on species' interaction patterns. Notably, the closest species in terms of foraging functional traits was the elephant *Loxodonta africana*, which also happen to be the closest evolutionary relative of the Hyracoidea in the Serengeti National Park.

The low-dimensional functional trait space of the Weddell Sea food web also exhibited a complex structure. In particular, the foraging functional traits distribution showed a split of the species into two well-defined groups. As we show below, this separation was strongly predicted by the phylogeny. In addition, it is related to the species' feeding behaviour and type. While the Serengeti food web showed a strong separation between plants and animals, the Weddell Sea trait space appeared to be more blurred across trophic guilds.

Phylogenetic signal

We tested for phylogenetic signal of species' functional roles estimated from the Seregenti and Weddell Sea food webs. In general, species' roles exhibited significant phylogenetic signal both for species as predators, as prey, or for both combined. Moreover, in both of these food webs, we observed a saturation effect in the correlation between the Jaccard similarity matrix and the phylogenetic variance–covariance matrix as we considered increasing model dimensions from $d = 1$ to $d = 20$ (Fig. 5). The pattern is even clearer, although less straightforward to read and may possibly depend on statistical artefacts, if we considered the contribution of single coordinates, as we have a decreasing signal (which is non significant for $d = 4$ and above in the Serengeti). The complete analysis outcome is presented in the Supplementary material.

Discussion

Previous research has indicated that simple, phenomenological food-web models can be successful but are unable to account for all the observed variance of food-web structure (Allesina et al. 2008, Rohr et al. 2010, Williams et al. 2010). To more accurately explain food-web structure, we therefore need to adopt a different or improved approach. Here, we introduce one such possibility, the directed random dot product graph model, and study its behavior for nine food webs (two larger ones from the Serengeti National Park and the Weddell Sea and seven smaller food webs). Having estimated the functional traits for the species in the food webs, we demonstrate that our model can fit observed interactions with considerable link-wise accuracy. We show also that the model can predict interactions for which we simulated absence of observation. While the enumeration and identification of the minimum or sufficient number of traits to 'explain' a food web is still an open problem (Eklöf et al. 2013, Capitán et al. 2013), our results support the argument that food webs are inherently low dimensional. In our approach, we distinguished between species' vulnerability and foraging traits with the former defining their 'role' as 'prey' and the latter their role as 'predators'.This distinction is not uncommon in the available literature (Bersier and Kehrli 2008, Rossberg et al. 2010, Rossberg 2013) and may help to explain an element of its success. Having said that, the way in which we identify these trait values is quite different to earlier approaches in the following way: previous research had attempted to define a suitable function F that maps a pair of vulnerability and foraging traits to the probability or occurrence of an interaction (Rossberg et al. 2006,

Rossberg 2013). Typically, the shape of F was proposed based on some natural assumption of species behaviour e.g. F should increase with respect to the similarity of the prey's vulnerability and predator's foraging traits (Petchey et al. 2008a) or F should allow for a certain dietary plasticity (Williams et al. 2010).

Actual traits involved in the process of food web assembly may differ across a web or for each pair of species. For instance, the traits determining the predation habits of an eagle may not be the same as those that determine the grazing preferences of a gazelle. Defining a suitable function F, many prior approaches had to face this problematic complexity imposing ad hoc variables and rules for each pair of species is a truly ambitious task, to say the least. Instead, we propose the use of abstract 'functional' traits that express the combined effect of many species-specific traits that would be measured empirically. Furthermore, our RDPG model adopts an extremely simple function F – a dot product – still able to satisfy key criteria (Rossberg 2013) while being remarkably predictive. A key consequence of this simplicity is that the complexity of explaining empirical interactions is shifted from identifying a suitable function F to the functional traits of the species themselves.

Similar approaches are not without precedent (Matias and Robin 2014). For example, the functional grouping of species proposed by Allesina and Pascual (2009) aims to identify groups of species that have distinct within- and between-group interaction probabilities. Within this group-focused framework – also known as stochastic block models (Holland et al. 1983, Wang and Wong 1987) – the species in a network are partitioned into groups (blocks) such that every group is non-empty and no node sits at the intersection between groups. As this clustering into groups is based solely on the observed food-web structure, the RDPG approach may be considered a generalization of stochastic block models that extends the upper bound on the number of blocks to the number of species in the community. Furthermore, if we consider a model where the block each species belongs to is determined by its position in 'functional space' (Rohe et al. 2011, Xu et al. 2014), we can recover the undirected RDPG model by assuming that the probability of a link between two nodes is given by their distance in this space. However, as we have seen comparing the fitting performance of our model and Allesina and Pascual's 2009 model, information is lost forcing the species into k blocks (groups) and assuming that species in blocks behave homogeneously.

Conclusions

Here, we have shown that complex food webs can be modeled with high fidelity based on a parsimonious stochastic model. Based on food webs' sequence of singular values, we consistently found a reasonable dimensionality upperbound much lower than the full rank of the food webs' adjacency matrix. The RDPG model outperformed other classic models both in terms of fitting performance – how many observed present/absent interactions were accurately captured – and in terms of predicting performance – how many non-observed interactions where accurately predicted. On this basis, we believe it important to distinguish the two performance frameworks and to explicitly consider unobserved interactions. We argue this may be best achieved by adopting a probabilistic view of species interactions.

Moreover, we detected a low but significant phylogenetic signal in the species' food-web roles – a result that echoes the conclusions of previous research (Bersier and Kehrli 2008, Stouffer et al. 2012, Rohr and Bascompte 2014). Here, however, we could distinguish between the contribution given by food webs' 'backbones' – the relative lower-dimensional model structure – and food webs' fine wiring – the relative higher-dimensional model structure. In particular, our results suggest that most of the evolutionary signal is already present in the structure of food webs' stochastic backbones. This pattern was consistently found when we considered independently species as consumers (or predators) and species as resources (or prey). Moreover, the saturating trend we detected when considering dimensionally increasing models was backed up by the analysis of single coordinates.

The fact that the predictive power of phylogeny varies as a function of the choice of model dimensionality begs the question of whether deterministic food-web models are really able to convey information about the evolutionary character of species-rich community. Confirming the presence of evolutionary signal in food webs, our results may be considered a first step in the direction of investigating more of the detailed nature of this signal. Again, we tenatively conclude that a probabilistic view of food webs may well represent a more suitable framework for such analysis.

Acknowledgements – All the code and data used in this paper is available on a public Github repository (http://gvdr.github.io/). We would like to thank Devin De Zwaan, Kendra Munn, Jenna Hutchen and Arne Mooers for help compiling the dated phylogenetic tree for the Serengeti food web, Carey Priebe and Minh Tanh for introducing us to RDPG models, Paul Brouwers for assisting us with the computational tasks, and Mike Steel, Jason Tylianakis, Timothée Poisot, Nick Baker, Ana-Johanna Voinopol-Sassu, Kate Wootton, Stinus Lindgreen, Melissa Broussard, Camille Coux, Carla Gomez-Creutzberg, Anuj Misra and Alyssa Cirtwill for helpful comments on an early draft. Funding to GVDR was provided by the Allan Wilson Centre for Molecular Ecology and Evolution and to DBS by a Marsden Fund Fast-Start grant (UOC-1101) and a Rutherford Discovery Fellowship, both administered by the Royal Society of New Zealand.

References

Allesina, S. 2011. Predicting trophic relations in ecological networks: a test of the allometric diet breadth model. – J. Theor. Biol. 279: 161–168.

Allesina, S. and Pascual, M. 2009. Food web models: a plea for groups. – Ecol. Lett. 12: 652–662.

Allesina, S. et al. 2008. A general model for food web structure. – Science 320: 658–661.

Andrews, H. C. and Patterson III, C. 1976. Singular value decomposition (SVD) image coding. – Communications, IEEE Trans. on 24: 425–432.

Baskerville, E. B. et al. 2011. Spatial guilds in the Serengeti food web revealed by a Bayesian group model. – PLoS Comput. Biol. 7: e1002321.

Bersier, L.-F. and Kehrli, P. 2008. The signature of phylogenetic constraints on food-web structure. – Ecol. Complex. 5: 132–139.

Black, A. J. and McKane, A. J. 2012. Stochastic formulation of ecological models and their applications. – Trends Ecol. Evol. 27: 337–345.

Canard, E. et al. 2014. Empirical evaluation of neutral interactions in host–parasite networks. – Am. Nat. 183: 468–479.
Capitán, J. A. et al. 2013. Degree of intervality of food webs: from body-size data to models. – J. Theor. Biol. 334: 35–44.
Cattell, R. B. 1966. The scree test for the number of factors. – Multivariate Behav. Res. 1: 245–276.
Chatterjee, S. et al. 2014. Matrix estimation by universal singular value thresholding. – Ann. Stat. 43: 177–214.
Closs, G. and Lake, P. 1994. Spatial and temporal variation in the structure of an intermittent stream food web. – Ecol. Monogr. 64: 2–21.
Cohen, J. E. 1977. Food webs and the dimensionality of trophic niche space. – Proc. Natl Acad. Sci. 74: 4533–4536.
Cohen, J. E. 1978. Food webs and niche space. No. 11 in Monographs in Population Biology. – Princeton Univ. Press.
Cohen, J. E. 1983. Recent progress and problems in food web theory. – Curr. Trends Food Web Theor. Oak Ridge Natl Lab. pp. 17–25.
Cohen, J. E. and Palka, Z. J. 1990. A stochastic theory of community food webs. V. Intervality and triangulation in the trophic-niche overlap graph. – Am. Nat. 135: 435–463.
Coulson, T. et al. 2004. Skeletons, noise and population growth: the end of an old debate? – Trends Ecol. Evol. 19: 359–364.
Dawah, H. A. et al. 1995. Structure of the parasitoid communities of grass-feeding chalcid wasps. – J. Anim. Ecol. 64: 708–720.
Eklöf, A. et al. 2013. The dimensionality of ecological networks. – Ecol. Lett. 16: 577–583.
Erdös, P. and Rényi, A. 1960. On the evolution of random graphs. – Publ. Math. Inst. Hung. Acad. Sci. 5: 17–61.
Felsenstein, J. 1985. Phylogenies and the comparative method. – Am. Nat. 125: 1–15.
Fortuna, M. A. et al. 2013. Habitat loss and the disassembly of mutalistic networks. – Oikos 122: 938–942.
Gavish, M. and Donoho, D. L. 2014. The optimal hard threshold for singular values is $4/\sqrt{3}$. – Information Theory, IEEE Trans. 60: 5040–5053.
Gravel, D. et al. 2013. Inferring food web structure rom predator–prey body size relationships. – Meth. Ecol. Evol. 4: 1083–1090.
Harper-Smith, S. et al. 2005. Communicating ecology through food webs: visualizing and quantifying the effects of stocking alpine lakes with trout. – In: De Ruiter, P. et al. (eds), Dynamic food webs: multispecies assemblages, ecosystem development and environmental change. Elsevier/Academic Press, pp. 407–423.
Holland, P. W. et al. 1983. Stochastic blockmodels: first steps. – Social Networks 5: 109–137.
Holling, C. S. 1973. Resilience and stability of ecological systems. – Annu. Rev. Ecol. Syst. 4: 1–23.
Jacob, U. et al. 2011. The role of body size in complex food webs: a cold case. – Adv. Ecol. Res. 45: 181–223.
Jennings, S. et al. 2002. Long-term trends in the trophic structure of the North Sea fish community: evidence from stable-isotope analysis, size-spectra and community metrics. – Mar. Biol. 141: 1085–1097.
Jonsson, T. et al. 2005. Food webs, body size, and species abundance in ecological community description. – Adv. Ecol. Res. 36: 1–84.
Jordán, F. 2009. Keystone species and food webs. – Phil. Trans. R. Soc. B 364: 1733–1741.
Lyzinski, V. et al. 2013. Perfect cluster- ing for stochastic blockmodel graphs via adjacency spectral embedding. – arXiv preprint arXiv:1310.0532 .
Matias, C. and Robin, S. 2014. Modeling heterogeneity in random graphs: a selective review. – arXiv:1402.4296 .
May, R. M. 2006. Network structure and the biology of populations. – Trends Ecol. Evol. 21: 394–399.
Memmott, J. et al. 2000. Predators, parasitoids and pathogens: species richness, trophic generality and body sizes in a natural food web. – J. Anim. Ecol. 69: 1–15.
Morowitz, H. J. 1980. The dimensionality of niche space. – J. Theor. Biol. 86: 259–263.
Mullon, C. et al. 2009. A minimal model of the variability of marine ecosystems. – Fish Fish. 10: 115–131.
Nickel, C. L. M. 2007. Random dot product graphs: a model for social networks. – PhD. thesis, The Johns Hopkins Univ.
Pagel, M. D. 1992. A method for the analysis of comparative data. – J. Theor. Biol. 156: 431–442.
Petchey, O. L. et al. 2008a. Size, foraging and food web structure. – Proc. Natl Acad. Sci. 105: 4191–4196.
Petchey, O. L. et al. 2008b. Trophically unique species are vulnerable to cascading extinction. – Am. Nat. 171: 568–579.
Poisot, T. et al. 2014. Beyond species: why ecological interactions vary through space and time. – bioRxiv.
Real, R. and Vargas, J. M. 1996. The probabilistic basis of Jaccard's index of similarity. – Syst. Biol. 45: 380–385.
Revell, L. J. et al. 2008. Phylogenetic signal, evolutionary process and rate. – Syst. Biol. 57: 591–601.
Rohe, K. et al. 2011. Spectral clustering and the high-dimensional stochastic blockmodel. – Ann. Stat. 39: 1878–1915.
Rohr, R. P. and Bascompte, J. 2014. Components of phylogenetic signal in antagonistic and mutualistic networks. – Am. Nat. 184: 556–564.
Rohr, R. P. et al. 2010. Modeling food webs: exploring unexplained structure using latent traits. – Am. Nat. 176: 170–177.
Rossberg, A. G. 2013. Food webs and biodiversity: foundations, models, data. – Wiley.
Rossberg, A. et al. 2006. Food webs: experts consuming families of experts. – J. Theor. Biol. 241: 552–563.
Rossberg, A. et al. 2010. How trophic interaction strength depends on traits. – Theor. Ecol. 3: 13–24.
Stouffer, D. B. 2010. Scaling from individuals to networks in food webs. – Funct. Ecol. 24: 44–51.
Stouffer, D. et al. 2005. Quantitative patterns in the structure of model and empirical food webs. – Ecology 86: 1301–1311.
Stouffer, D. B. et al. 2006. A robust measure of food web intervality. – Proc. Natl Acad. Sci. 103: 19015–19020.
Stouffer, D. B. et al. 2011. The role of body mass in diet contiguity and food-web structure. – J. Anim. Ecol. 80: 632–639.
Stouffer, D. B. et al. 2012. Evolutionary conservation of species' roles in food webs. – Science 335: 1489–1492.
Tang, M. et al. 2013. Universally consistent vertex classification for latent positions graphs. – Ann. Stat. 41: 1406–1430.
Wang, Y. J. and Wong, G. Y. 1987. Stochastic blockmodels for directed graphs. – J. Am. Stat. Ass. 82: 8–19.
Wasserman, S. 1994. Social network analysis: methods and applications. Vol. 8. – Cambridge Univ. Press.
Williams, R. J. and Martinez, N. D. 2000. Simple rules yield complex food webs. – Nature 404: 180–183.
Williams, R. J. and Martinez, N. D. 2008. Success and its limits among structural models of complex food webs. – J. Anim. Ecol. 77: 512–519.
Williams, R. J. et al. 2010. The probabilistic niche model reveals the niche structure and role of body size in a complex food web. – PloS ONE 5(8): e12092.
Woodward, G. and Hildrew, A. G. 2001. Invasion of a stream food web by a new top predator. – J. Anim. Ecol. 70: 273–288.
Woodward, G. et al. 2005a. Body size in ecological networks. – Trends Ecol. Evol. 20: 402–409.

Woodward, G. et al. 2005b. Quantification and resolution of a complex, size-structured food web. – Adv. Ecol. Res. 36: 85–135.
Xu, J. et al. 2014. Edge label inference in generalized stochastic block models: from spectral theory to impossibility results. – arXiv:1406.6897.
Young, S. J. and Scheinerman, E. R. 2007. Random dot product graph models for social networks. – In: Algorithms and models for the web-graph. Springer, pp. 138–149.
Young, S. J. and Scheinerman, E. R. 2008. Directed random dot product graphs. – Internet Math. 5: 91–111.
Zhu, M. and Ghodsi, A. 2006. Automatic dimensionality selection from the scree plot via the use of profile likelihood. – Comput. Stat. Data Anal. 51: 918–930.
Zook, A. E. et al. 2011. Food webs: ordering species according to body size yields high degree of intervality. – J. Theor. Biol. 271: 106–113.

Supplementary material (available online as Appendix oik.02305 at <www.oikosjournal.org/readers/appendix>).
Appendix 1.