

# Food web reconstruction through phylogenetic transfer of low-rank network representation

Tanya Strydom<sup>1,2,‡</sup>, Salomé Bouskila<sup>1,2,‡</sup>, Timothée Poisot<sup>1,2</sup>

<sup>1</sup> Université de Montréal; <sup>2</sup> Québec Centre for Biodiversity Sciences

‡ These authors contributed equally to the work

## Correspondance to:

Timothée Poisot — timothee.poisot@umontreal.ca

---

**Abstract:** TODO

**Keywords:**

ecological networks

network embedding

transfer learning

ancestral character estimation

biogeography

---

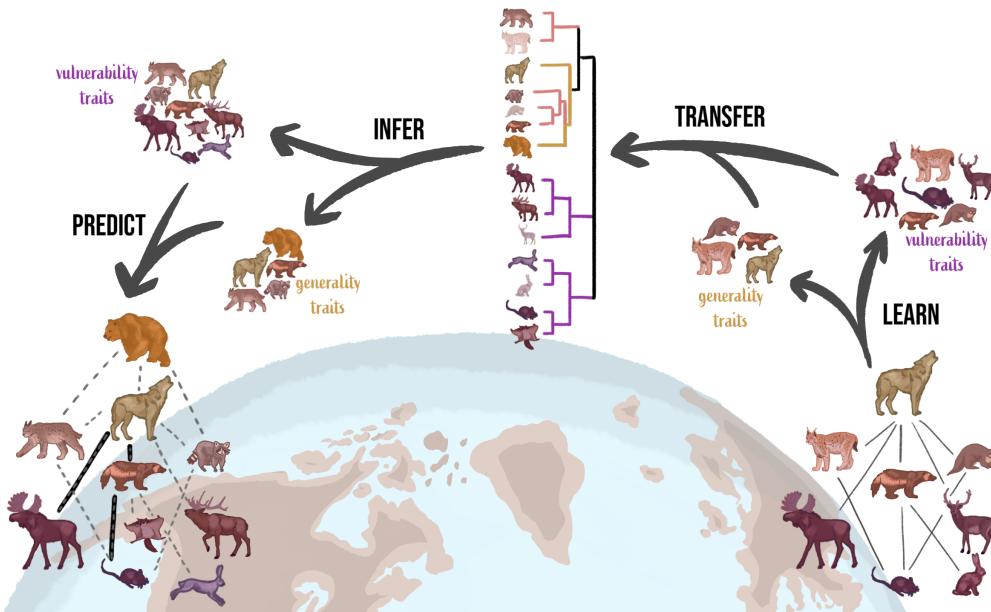
1

## Introduction

There are two core challenges to our understanding of ecological networks in space, particularly at macro-ecologically relevant scales (*e.g.* Trojelsgaard and Olesen 2016). First, networks within a location are difficult to sample properly (Jordano 2016a, 2016b), and this results in a widespread “Eltonian shortfall” (Hortal et al. 2015). This first challenge (local incompleteness) has been in large part addressed by the recent multiplication of methods aiming to predict interactions within an existing network, a lot of which are reviewed in Strydom et al. (2021). Second, recent analyses based on collected data (Poisot, Bergeron, et al. 2021) or metadata (Cameron et al. 2019) highlight that ecological networks are currently studied in a biased subset of space and bioclimates, which impedes our ability to generalize any local understanding of network structure. Having a general solution for the issue of metaweb inference (Morales-Castilla et al. 2015) even in situations where minimal knowledge about existing interactions within the species pool is known would catalyze significant breakthroughs in our ability to start thinking about species interactions networks over large spatial scales.

Here, we present a general method for the transfer learning of network representations, relying on similarities of species in a biologically relevant proxy space. This method allows to the prediction of species interactions based on knowledge extracted from another species pool. Our method is outlined in fig. 1; we provide an illustration based on learning an embedding of a metaweb of trophic interactions for European mammals (Maiorano et al. 2020b, 2020a), and based on phylogenetic relationships between mammals globally (Upham, Esselstyn, and Jetz 2019), transfer this representation to the pool of mammals in Canada.

There are a plurality of measures of species similarities that can be used for knowledge transfer (see *e.g.* Morales-Castilla et al. 2015); nevertheless, phylogenetic proximity has a number of desirable properties when working at large scales. Gerhold et al. (2015) make the point that



**Figure 1** Overview of the phylogenetic transfer learning of species interactions networks. Starting from an initial network, we learn its representation through a graph embedding step (here, a truncated Singular Value Decomposition), yielding a series of latent traits; second, for the destination species pool, we perform ancestral character estimation using a phylogeny (here, using a Brownian model for the latent traits); we then sample from the reconstructed distribution of latent traits to generate a probabilistic metaweb at the destination (here, assuming a uniform distribution of traits), and threshold it to yield the final list of interactions.

phylogenetic signal captures diversification of characters (large macro-evolutionary process), but not necessarily community assembly (fine ecological process); Dormann et al. (2010) previously gave very similar conclusions. Interactions tend to conserve phylogenetic signal that encompasses a wide range of ecological and evolutionary mechanisms (Mouquet et al. 2012; Cavender-Bares et al. 2009), and most importantly, retain this signal even when it is not detectable at the community scale (Poisot and Stouffer 2018; Hutchinson, Cagua, and Stouffer 2017). Finally, species interactions at macro-ecological scales seem to respond mostly to macro-evolutionary processes (Price 2003); this is evidenced by the presence of conserved backbones in food webs (Dalla Riva and Stouffer 2016), strong evolutionary signature on prey choice (Stouffer et al. 2012), and strong phylogenetic signature in food web intervalty (Eklöf and Stouffer 2016). Taken together, these consideration suggest that phylogenies can be reliably be used to transfer knowledge on species interactions.

TK Results/summary type paragraph

## 2

### Data used for the case study

We use data on the European metaweb assembled by Maiorano et al. (2020b). We use the definition of the metaweb first introduced by Dunne (2006), *i.e.* an inventory of all possible interactions within species from a spatially delimited pool. The metaweb is not a prediction of the food web at any specific locale within the frontiers of the species pool – in fact, these local food webs are expected to have a subset of both the species and the interactions of their metaweb (Poisot et al. 2012). This being said, as the metaweb represents the total of functional, phylogenetic, and macroecological processes (Morales-Castilla et al. 2015), it is still worthy of ecological attention. We induced the subgraph corresponding to all mammals by matching species names in the original network first to the GBIF taxonomic backbone (GBIF Secretariat 2021), and retaining all those who matched to mammals; all nodes had valid matches to GBIF at this step, and so this backbone is used for all name reconciliation steps as outlined below.

The European metaweb represents the knowledge we want to learn and transfer; the support for transfer is here represented by the phylogenetic similarity of mammals. We used the mammalian

consensus supertree by Upham, Esselstyn, and Jetz (2019), for which all approx. 6000 names have been similarly matched to their GBIF valid names. This step allows us to place each node of the mammalian European metaweb in the phylogeny.

The destination problem to which we want to transfer knowledge is the trophic interactions between mammals in Canada. We obtained the list of extant species from the IUCN checklist, and selected the terrestrial and semi-aquatic species (this corresponds to the same selection that was applied by Maiorano et al. (2020b) in the European metaweb). The IUCN names were, as previously, reconciled against GBIF to have an exact match to the taxonomy.

The European metaweb had XXX species, and the Canadian species pool had YYY - of these, only 17 were shared, which highlights the need to rely on a knowledge transfer approach. In the following sections, we describe the representational learning step applied to European data, the transfer step through phylogenetic similarity, and the generation of a probabilistic metaweb for the destination species pool.

## 3

---

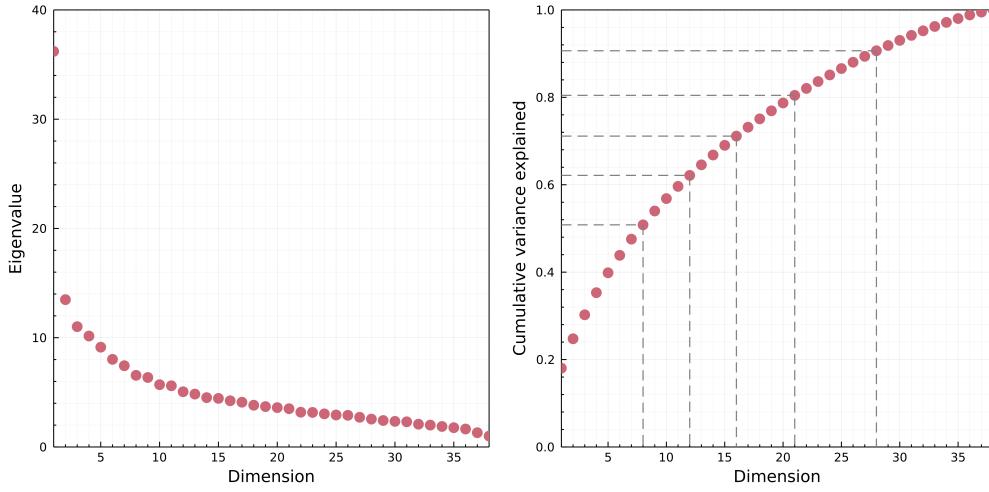
### Method description

The crux of the method is the transfer of knowledge of a known network, in order to predict interactions between species from another location. In fig. 1, we give a high-level overview of the approach; in the example around which this manuscript is built (leveraging detailed knowledge about trophic interactions between Mammalia in Europe to predict the less known trophic interactions between closely phylogenetically related Mammalia in Canada), we use a series of specific steps for network embedding, trait inference, and network prediction and thresholding. This being said, the methods at each step may be substituted by others when the data support it.

Specifically, our approach can be summarized as follows: from the known network in Europe, we use a truncated Singular Values Decomposition (t-SVD) to generate latent traits representing a low-dimensional embedding of the network; these traits give an unbiased estimate of the node's position in the latent feature spaces. Then, we map these latent traits onto a reference phylogeny (other distance-based measures of species proximity that allow for the inference of features in the latent space can be used, such as for example dissimilarity in functional traits). Based on the reconstructed latent traits for species in the destination species pool, a Random Dot Product Graph (RDPG) model predicts the interaction between species through a function of the nodes' features. Thus, from latent traits and nodes position, we can infer interactions.

**3.1. Learning the origin network representation** The first step in transfer learning is to learn the structure of the origin dataset. In order to do so, we rely on an approach inspired from representational learning, where we learn a *representation* of the metaweb, rather than a list of interactions. This approach is conceptually different from other metaweb-scale predictions (*e.g.* ALBOUY, others?), in that the metaweb representation is easily transferable. Specifically, we use Random Dot Product Graphs (RDPG; Young and Scheinerman 2007) to create a number of latent variables that can be combined into an approximation of the network adjacency matrix. RDPG results are known to have strong phylogenetic signal, and to capture the evolutionary backbone of food webs (Dalla Riva and Stouffer 2016). In addition, recent advances show that the latent variables produced this way can be used to predict *de novo* network edges (Runghen, Stouffer, and Dalla Riva 2021).

The latent variables are created by performing a truncated Singular Value Decomposition (tSVD) on the adjacency matrix. SVD is an appropriate embedding of ecological networks, which has recently been shown to both capture their complex, emerging properties (Strydom, Dalla Riva, and Poisot 2021), and to allow highly accurate prediction of the interactions within a single



**Figure 2** Left: representation of the screeplot of the eigenvalues from the tSVD on the European metaweb. The screeplot shows no obvious drop in the eigenvalues that may be leveraged to automatically detect a minimal dimension for embedding, after e.g. Zhu and Ghodsi (2006). Right: cumulative fraction of variance explained by each dimensions up to the rank of the European metaweb. The grey lines represents cutoff at 40, 60, and 80% of variance explained. For the rest of the analysis, we reverted to an arbitrary threshold of 60% of variance explained, which represented a good tradeoff between accuracy and reduced number of features.

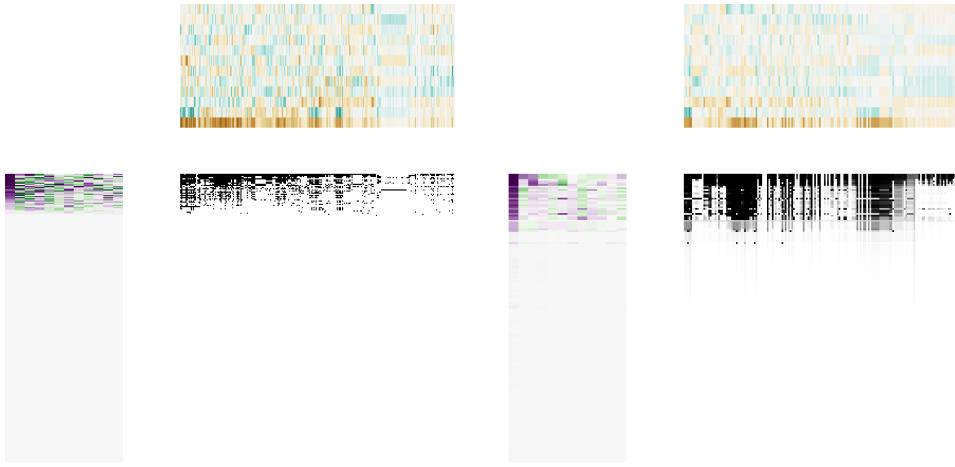
network (Poisot, Ouellet, et al. 2021). Under tSVD, an adjacency matrix  $\mathbf{A}$  is decomposed into three components, so that  $\mathbf{A} = \mathbf{L}\Sigma\mathbf{R}^*$ , for which  $\Sigma$  is an  $r \times r$  diagonal matrix (where  $r$  is the rank of matrix  $\mathbf{A}$ ) containing only non-zero singular ( $\sigma$ ) values,  $\mathbf{L}$  is an  $m \times r$  matrix, and  $\mathbf{R}$  an  $n \times r$  matrix. In using a tSVD we retain only ‘informative’ (non-zero)  $\sigma$  values - which translates to unique predation strategies within the matrix.

The latent variables used for the RDPG, called the left and right subspaces, are defined as  $\mathcal{L} = \mathbf{L}\sqrt{\Sigma}$ , and  $\mathcal{R} = \sqrt{\Sigma}\mathbf{R}$  – using the full rank of  $\mathbf{A}$ ,  $\mathcal{L}\mathcal{R}' = \mathbf{A}$ , and using any smaller rank results in  $\mathcal{L}\mathcal{R}' \approx \mathbf{A}$ . Using a rank of 1 for the tSVD provides a first-order approximation of the network, etc.

The specific rank at which the SVD ought to be truncated is a difficult question. The purpose of SVD is to remove the noise, expressed at high dimensions, and to focus on the signal, expressed at low dimension. In datasets with a clear signal/noise demarcation, the scree plot of  $\Sigma$  can show a sharp drop at the rank where noise starts. Alternatively, REF showed that the profile likelihood (assuming a known distribution of  $\Sigma$ , or a normal distribution otherwise) peaks at the rank maximizing the signal-to-noise ratio. Because the European metaweb is almost entirely known, the amount of noise is low; this is reflected in fig. 2 (left), where the scree plot shows no important drop, and in fig. 2 (right) where the proportion of variance explained increases smoothly at higher dimensions. For this reason, we default back to an arbitrary threshold that explain 60% of the variance in the underlying data, corresponding to 12 dimensions.

Because RDPG relies on matrix multiplication, the higher dimensions essentially serve to make specific interactions converge towards 0 or 1; therefore, for reasonably low ranks, there is no guarantee that the values in the reconstructed network will be within the unit range. In order to determine what constitutes an appropriate threshold, we performed the RDPG approach on the European metaweb, and evaluated the threshold by treating this as a binary classification problem, specifically assuming that both 0 and 1 in the European metaweb are all true. Given the methodological given in REF, this seems like a reasonable assumption, although one that does not hold for all metawebs. We used the thresholding approach presented in HAYSTACK, and picked a cutoff that maximized Youden’s J statistic (Youden 1950); the resulting cutoff was 0.22, and gave an accuracy above 0.99.

The left and right subspaces for the European metaweb, accompanied by the threshold, represent the knowledge we seek to transfer. In the next section, we explain how we rely on phylogenetic similarity to do so.



**Figure 3** Visual representation of the left (green/purple) and right (green/brown) subspaces, alongside the adjacency matrix of the food web they encode (greyscale). The European metaweb is on the left, and the imputed Canadian metaweb (before data inflation) on the right. This figure illustrates how much structure the left sub-space captures. As we show in fig. 6, the species with a value of 0 in the left subspace are species without any preys.

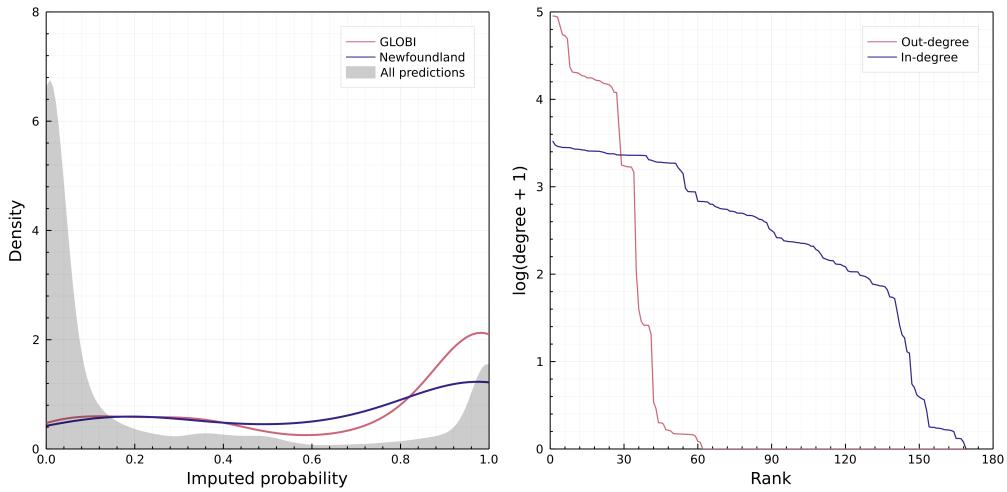
**3.2. Transfer learning through phylogenetic relatedness** In order to transfer the knowledge from the European metaweb to the Canadian species pool, we performed ancestral character estimation using a Brownian motion model, which is a conservative approach in the absence of strong hypotheses about the nature of phylogenetic signal in the network decomposition (Litsios and Salamin 2012). We assumed that all traits (*i.e.* the feature vectors for the left and right subspaces) were independent, which is a reasonable assumption as every trait/dimension added to the tSVD has an *additive* effect to the one before it. The Brownian motion algorithm returns the *average* value of the trait, and its upper and lower bounds. Because we do not estimate other parameters of the traits distributions, we considered that every species trait is represented as a uniform distribution between these bounds. Therefore, the inferred left and right subspaces for the Canadian species pool ( $\hat{\mathcal{L}}$  and  $\hat{\mathcal{R}}$ ) have entries that are distributions, representing the range of values for a given species at a given dimension.

These objects represent the transferred knowledge, which we can use for prediction of the Canadian metaweb.

**3.3. Probabilistic prediction of the destination network** The phylogenetic reconstruction of  $\hat{\mathcal{L}}$  and  $\hat{\mathcal{R}}$  has an associated uncertainty, represented by the breadth of the uniform distribution associated to each of their entries. Therefore, we can use this information to assemble a *probabilistic* metaweb in the sense of Poisot et al. (2016), *i.e.* in which every interaction is represented as a single, independent, Bernoulli event of probability  $p$ .

Specifically, we have adopted the following approach. For every entry in  $\hat{\mathcal{L}}$  and  $\hat{\mathcal{R}}$ , we draw a value from its distribution. This results in one instance of the possible left () and right () subspaces for the Canadian metaweb. These can be multiplied, to produce one matrix of real values. Because the entries in  $\hat{\mathcal{L}}$  and  $\hat{\mathcal{R}}$  are in the same space where  $\mathcal{L}$  and  $\mathcal{R}$  were originally predicted, it follows that the threshold  $\rho$  estimated for the European metaweb also applies. We use this information to produce one random Canadian metaweb,  $N = \hat{\mathcal{L}}\hat{\mathcal{R}}' \geq \rho$  fig. 3.

Because the intervals around some trait values can be broad (in fact, probably broader than what they would actually be, see *e.g.* Garland, Midford, and Ives 1999), we repeat the above process  $2 \times 10^5$  times, which results in a probabilistic metaweb  $P$ , where the probability of an interaction (here conveying our degree of trust that it exists given the inferred trait distributions) is given by the number of times where it appears across all random draws  $N$ , divided by the number of samples. An interaction with  $P_{i,j} = 1$  means that these two species were predicted to interact in all  $2 \times 10^5$  random draws, etc..



**Figure 4** GLOBI comparison REF TK/TP

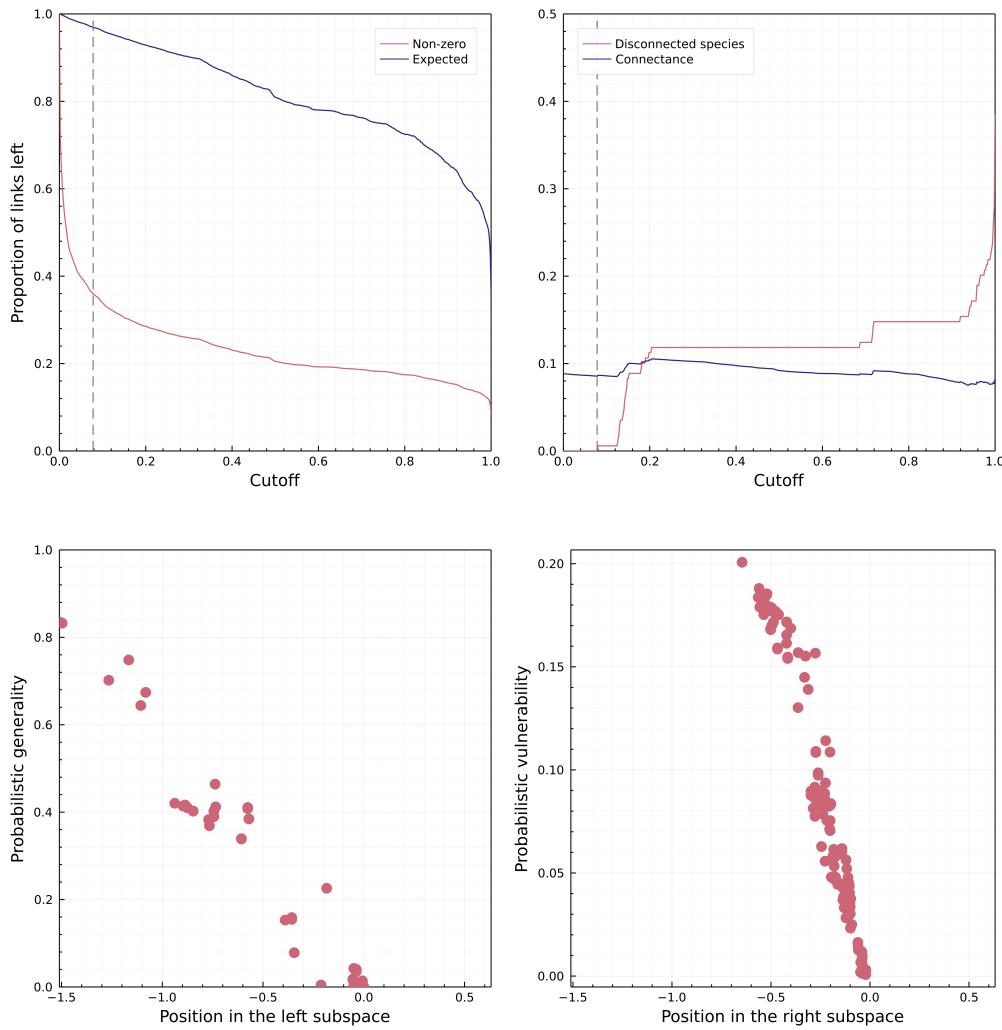
**3.4. Data cleanup, discovery, validation, and thresholding** Once the probabilistic metaweb for Canada has been produced, we followed a number of data inflation steps to finalize it.

First, we extracted the subgraph corresponding to the 17 species shared between the European and Canadian pools, and replaced these interactions with a probability of 0 (non-interaction) or 1 (interaction). This represents a minute modification of the inferred network (about 0.8% of all species pairs from the Canadian web), but ensures that are directly re-using knowledge from Europe.

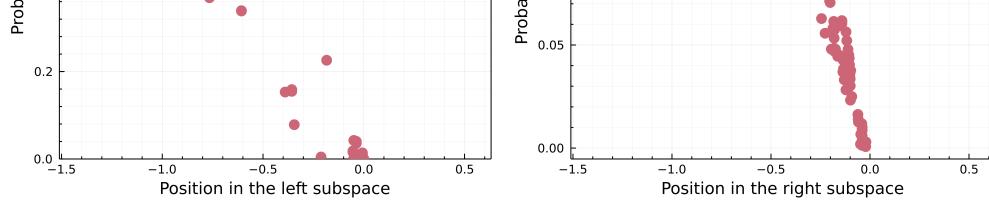
Second, we looked for all species in the Canadian pool known to the Global Biotic Interactions (GLOBI) database **REF**, and extracted their known interactions. Because GLOBI aggregates observed interactions, it is not a *networks* data source, and therefore the only information we can reliably extract from it is that a species pair *was reported to interact at least once*. This last statement should yet be taken with caution, as some sources in GLOBI (*e.g.* **TK**) are produced through text-mining, and therefore do not document direct evidence of the interaction. Nevertheless, should the predictive model work, we would expect that a majority of interactions known to GLOBI would also be predicted. After performing this check, we set the probability of all interactions known to GLOBI (366 in total, 33 of which were not predicted by the model, for a success rate of 91%) to 1.

Finally, we downloaded the data from Strong and Leroux (2014), who mined various literature sources to identify trophic interactions in Newfoundland. This dataset had documented 25 interactions between mammals, only two of which were not part of our predictions, for a success rate of 92%. These two interactions were added to our predicted metaweb, with a probability of 1.

Because the confidence interval on the inferred trait space are probably an over-estimate, we decided to apply a thresholding step to the interactions after the data inflation fig. 5. Cirtwill and Hambäck (2021) highlight a number of strategies to threshold probabilistic networks. Their methods assume the underlying data to be tag-based sequencing, which represents interactions as co-occurrences of predator and prey within the same tags; this is conceptually identical to our Bernoulli-trial based reconstruction of a probabilistic network. We performed a full analysis of the effect of various cutoffs, and as they either resulted in removing to few interactions, or removing enough interactions that species started to be disconnected from the network, we set the threshold for a probability equivalent to 0 to the largest possible value that still allow all species to have at least one interaction with non-zero probability.



**Figure 5** Left: effect of varying the cutoff for probabilities to be considered non-zero on the number of unique links and on  $\hat{L}$ , the probabilistic estimate of the number of links assuming that all interactions are independant. Right: effect of varying the cutoff on the number of disconnected species, and on network connectance. In both panels, the grey line indicates the cutoff  $\rho = TK$  that resulted in the first species losing all of its interactions.



**Figure 6** Left: there is a linear relationship between the values on the first dimension of the left subspace and the generality, *i.e.* the relative number of preys, *sensu* Schoener (1989). Species with a value of 0 in this subspace are at the bottom-most trophic level. Right: there is, similarly, a linear relationship between the position of a species on the first dimension of the right subspace and its vulnerability, *i.e.* the relative number of predators. Taken together, these two figures show that the first-order representation of this network would capture its degree distribution.

**3.5. Implementation and code availability** The entire pipeline was implemented in *Julia* 1.6 (Bezanson et al. 2017), and is available under the permissive MIT License at DOI. The taxonomic cleanup steps are done using GBIF.jl (Dansereau and Poisot 2021). The network embedding and analysis is done using EcologicalNetworks.jl (Banville, Vissault, and Poisot 2021; Poisot et al. 2019). The phylogenetic simulations are done using PhyloNetworks.jl (Solís-Lemus, Bastide, and Ané 2017) and Phylo.jl (Reeve et al. 2016). A complete Project.toml file specifying the full tree of dependencies is available alongside the code, at <https://osf.io/2zwqm/>.

## 4

## Results and discussion of the case study

TP fig. 5 – the cutoffs using maximum curvature of central difference approximation of the second order partial derivative result in respectively species being lost, or almost all links being kept – we settled on the value that allowed all species to remain.

The t-SVD embedding is able to learn relevant ecological features for the network. fig. 6 shows that the first rank correlates linearly with generality and vulnerability (Schoener 1989), *i.e.* the

number of preys and predators. Importantly, this implies that a rank 1 approximation represents the configuration model (Park and Newman 2004) for the metaweb, or (accounting for the probabilistic nature of the degree) the soft configuration model (van der Hoorn, Lippner, and Krioukov 2018), both of which are maximum entropy graph models (Garlaschelli, Hollander, and Roccaverde 2018), with respectively sharp and soft local constraints. The configuration model is widely used by ecologists in the context of null hypothesis significance testing of network structure (*e.g.* Bascompte et al. 2003), and it is noteworthy that for this metaweb, the relevant information was extracted at the first rank. Because the first rank corresponds to the leading eigenvalue of the system, the results of fig. 6 have a straightforward interpretation: degree-based processes are the most important in structuring the mammalian food web.

---

## 5

## Conclusions

A cautionary tale regarding validation. Interaction can never be a true negative. Cannot use a empirical subset of a bigger (predicated) network for validation -> because *e.g.* dietary shifts across range.

Cirtwill et al. (2019) – we need an informative interaction-level prior on probability: this method might be it

---

## References

- Banville, Francis, Steve Vissault, and Timothée Poisot. 2021. “Mangal.jl and EcologicalNetworks.jl: Two Complementary Packages for Analyzing Ecological Networks in Julia.” *Journal of Open Source Software* 6 (61): 2721. <https://doi.org/10.21105/joss.02721>.
- Bascompte, J., P. Jordano, C. J. Melian, and J. M. Olesen. 2003. “The Nested Assembly of Plant-Animal Mutualistic Networks.” *Proceedings of the National Academy of Sciences* 100 (16): 9383–87. <https://doi.org/10.1073/pnas.1633576100>.
- Bezanson, J., A. Edelman, S. Karpinski, and V. Shah. 2017. “Julia: A Fresh Approach to Numerical Computing.” *SIAM Review* 59 (1): 65–98. <https://doi.org/10.1137/141000671>.
- Cameron, Erin K., Maja K. Sundqvist, Sally A. Keith, Paul J. CaraDonna, Erik A. Mousling, Karin A. Nilsson, Daniel B. Metcalfé, and Aimée T. Classen. 2019. “Uneven Global Distribution of Food Web Studies Under Climate Change.” *Ecosphere* 10 (3): e02645. <https://doi.org/10.1002/ecs2.2645>.
- Cavender-Bares, Jeannine, Kenneth H. Kozak, Paul V. A. Fine, and Steven W. Kembel. 2009. “The Merging of Community Ecology and Phylogenetic Biology.” *Ecology Letters* 12 (7): 693–715. <https://doi.org/10.1111/j.1461-0248.2009.01314.x>.
- Cirtwill, Alyssa R., Anna Eklf, Tomas Roslin, Kate Wootton, and Dominique Gravel. 2019. “A Quantitative Framework for Investigating the Reliability of Empirical Network Construction.” *Methods in Ecology and Evolution* 0 (ja). <https://doi.org/10.1111/2041-210X.13180>.
- Cirtwill, Alyssa R., and Peter Hambäck. 2021. “Building Food Networks from Molecular Data: Bayesian or Fixed-Number Thresholds for Including Links.” *Basic and Applied Ecology* 50 (February): 67–76. <https://doi.org/10.1016/j.baae.2020.11.007>.

- Dalla Riva, Giulio V, and Daniel B. Stouffer. 2016. “Exploring the Evolutionary Signature of Food Webs’ Backbones Using Functional Traits.” *Oikos* 125 (4): 446–56. <https://doi.org/10.1111/oik.02305>.
- Dansereau, Gabriel, and Timothée Poisot. 2021. “SimpleSDMLayers.jl and GBIF.jl: A Framework for Species Distribution Modeling in Julia.” *Journal of Open Source Software* 6 (57): 2872. <https://doi.org/10.21105/joss.02872>.
- Dormann, Carsten F., Bernd Gruber, Marten Winter, and Dirk Herrmann. 2010. “Evolution of Climate Niches in European Mammals?” *Biology Letters* 6 (2): 229–32. <https://doi.org/10.1098/rsbl.2009.0688>.
- Dunne, Jennifer A. 2006. “The Network Structure of Food Webs.” In *Ecological Networks: Linking Structure and Dynamics*, edited by Jennifer A Dunne and Mercedes Pascual, 27–86. Oxford University Press.
- Eklöf, Anna, and Daniel B. Stouffer. 2016. “The Phylogenetic Component of Food Web Structure and Intervality.” *Theoretical Ecology* 9 (1): 107–15. <https://doi.org/10.1007/s12080-015-0273-9>.
- Garland, THEODORE, JR., PETER E. Midford, and ANTHONY R. Ives. 1999. “An Introduction to Phylogenetically Based Statistical Methods, with a New Method for Confidence Intervals on Ancestral Values1.” *American Zoologist* 39 (2): 374–88. <https://doi.org/10.1093/icb/39.2.374>.
- Garlaschelli, Diego, Frank den Hollander, and Andrea Roccaverde. 2018. “Covariance Structure Behind Breaking of Ensemble Equivalence in Random Graphs.” *Journal of Statistical Physics* 173 (3-4): 644–62. <https://doi.org/10.1007/s10955-018-2114-x>.
- GBIF Secretariat. 2021. “GBIF Backbone Taxonomy.” <https://doi.org/10.15468/39omei>.
- Gerhold, Pille, James F. Cahill, Marten Winter, Igor V. Bartish, and Andreas Prinzing. 2015. “Phylogenetic Patterns Are Not Proxies of Community Assembly Mechanisms (they Are Far Better).” *Functional Ecology* 29 (5): 600–614. <https://doi.org/10.1111/1365-2435.12425>.
- Hoorn, Pim van der, Gabor Lippner, and Dmitri Krioukov. 2018. “Sparse Maximum-Entropy Random Graphs with a Given Power-Law Degree Distribution.” *Journal of Statistical Physics* 173 (3-4): 806–44. <https://doi.org/10.1007/s10955-017-1887-7>.
- Hortal, Joaquín, Francesco de Bello, José Alexandre F. Diniz-Filho, Thomas M. Lewinsohn, Jorge M. Lobo, and Richard J. Ladle. 2015. “Seven Shortfalls That Beset Large-Scale Knowledge of Biodiversity.” *Annual Review of Ecology, Evolution, and Systematics* 46 (1): 523–49. <https://doi.org/10.1146/annurev-ecolsys-112414-054400>.
- Hutchinson, Matthew C., Edgar Fernando Cagua, and Daniel B. Stouffer. 2017. “Cophylogenetic Signal Is Detectable in Pollination Interactions Across Ecological Scales.” *Ecology*, n/a-. <https://doi.org/10.1002/ecy.1955>.
- Jordano, Pedro. 2016a. “Chasing Ecological Interactions.” *PLOS Biology* 14 (9): e1002559. <https://doi.org/10.1371/journal.pbio.1002559>.
- . 2016b. “Sampling Networks of Ecological Interactions.” *Functional Ecology*, September. <https://doi.org/10.1111/1365-2435.12763>.
- Litsios, Glenn, and Nicolas Salamin. 2012. “Effects of Phylogenetic Signal on Ancestral State Reconstruction.” *Systematic Biology* 61 (3): 533–38. <https://doi.org/10.1093/sysbio/syr124>.

- Maiorano, Luigi, Alessandro Montemaggioli, Gentile Francesco Ficetola, Louise O'Connor, and Wilfried Thuiller. 2020a. "Data from: Tetra-EU 1.0: A Species-Level Trophic Meta-Web of European Tetrapods." Dryad. <https://doi.org/10.5061/DRYAD.JM63XSJ7B>.
- . 2020b. "TETRA-EU 1.0: A Species-Level Trophic Metaweb of European Tetrapods." Edited by Allen Hurlbert. *Global Ecology and Biogeography* 29 (9): 1452–57. <https://doi.org/10.1111/geb.13138>.
- Morales-Castilla, Ignacio, Miguel G. Matias, Dominique Gravel, and Miguel B. Araújo. 2015. "Inferring Biotic Interactions from Proxies." *Trends in Ecology & Evolution* 30 (6): 347–56. <https://doi.org/10.1016/j.tree.2015.03.014>.
- Mouquet, Nicolas, Vincent Devictor, Christine N. Meynard, Francois Munoz, Louis-Félix Bersier, Jérôme Chave, Pierre Couteron, et al. 2012. "Ecophylogenetics: Advances and Perspectives." *Biological Reviews* 87 (4): 769–85. <https://doi.org/10.1111/j.1469-185X.2012.00224.x>.
- Park, Juyong, and M. E. J. Newman. 2004. "Statistical Mechanics of Networks." *Physical Review E* 70 (6): 066117. <https://doi.org/10.1103/PhysRevE.70.066117>.
- Poisot, Timothée, Zacharie Belisle, Laura Hoebelke, Michiel Stock, and Piotr Szefer. 2019. "EcologicalNetworks.jl - Analysing Ecological Networks." *Ecography*. <https://doi.org/10.1111/ecog.04310>.
- Poisot, Timothée, Gabriel Bergeron, Kevin Cazelles, Tad Dallas, Dominique Gravel, Andrew MacDonald, Benjamin Mercier, Clément Violet, and Steve Vissault. 2021. "Global Knowledge Gaps in Species Interaction Networks Data." *Journal of Biogeography* n/a (n/a). <https://doi.org/10.1111/jbi.14127>.
- Poisot, Timothée, Elsa Canard, David Mouillot, Nicolas Mouquet, and Dominique Gravel. 2012. "The Dissimilarity of Species Interaction Networks." *Ecology Letters* 15 (12): 1353–61. <https://doi.org/10.1111/ele.12002>.
- Poisot, Timothée, Alyssa R. Cirtwill, Kévin Cazelles, Dominique Gravel, Marie-Josée Fortin, and Daniel B. Stouffer. 2016. "The Structure of Probabilistic Networks." Edited by Jana Vamosi. *Methods in Ecology and Evolution* 7 (3): 303–12. <https://doi.org/10.1111/2041-210X.12468>.
- Poisot, Timothée, Marie-Andrée Ouellet, Nardus Mollentze, Maxwell J. Farrell, Daniel J. Becker, Gregory F. Albery, Rory J. Gibb, Stephanie N. Seifert, and Colin J. Carlson. 2021. "Imputing the Mammalian Virome with Linear Filtering and Singular Value Decomposition." *arXiv:2105.14973 [q-Bio]*, May. <http://arxiv.org/abs/2105.14973>.
- Poisot, Timothée, and Daniel B. Stouffer. 2018. "Interactions Retain the Co-Phylogenetic Matching That Communities Lost." *Oikos* 127 (2): 230–38. <https://doi.org/10.1111/oik.03788>.
- Price, Peter W. 2003. *Macroevolutionary Theory on Macroecological Patterns*. Cambridge University Press.
- Reeve, Richard, Tom Leinster, Christina A. Cobbold, Jill Thompson, Neil Brummitt, Sonia N. Mitchell, and Louise Matthews. 2016. "How to Partition Diversity." *arXiv:1404.6520 [q-Bio]*, December. <http://arxiv.org/abs/1404.6520>.
- Runghen, Rogini, Daniel B Stouffer, and Giulio V Dalla Riva. 2021. "Exploiting Node Metadata to Predict Interactions in Large Networks Using Graph Embedding and Neural Networks," June. <https://doi.org/10.1101/2021.06.10.447991>.
- Schoener, Thomas W. 1989. "Food Webs from the Small to the Large." *Ecology* 70 (6): 1559–89.

- Solís-Lemus, Claudia, Paul Bastide, and Cécile Ané. 2017. “PhyloNetworks: A Package for Phylogenetic Networks.” *Molecular Biology and Evolution* 34 (12): 3292–98. <https://doi.org/10.1093/molbev/msx235>.
- Stouffer, D. B., M. Sales-Pardo, M. I. Sirer, and J. Bascompte. 2012. “Evolutionary Conservation of Species’ Roles in Food Webs.” *Science* 335 (6075): 1489–92. <https://doi.org/10.1126/science.1216556>.
- Strong, Justin S., and Shawn J. Leroux. 2014. “Impact of Non-Native Terrestrial Mammals on the Structure of the Terrestrial Mammal Food Web of Newfoundland, Canada.” *PLOS ONE* 9 (8): e106264. <https://doi.org/10.1371/journal.pone.0106264>.
- Strydom, Tanya, Michael David Catchen, Francis Banville, Dominique Caron, Gabriel Dansereau, Philippe Desjardins-Proulx, Norma Forero, et al. 2021. “A Roadmap Toward Predicting Species Interaction Networks (Across Space and Time).” Preprint. EcoEvoRxiv. <https://doi.org/10.32942/osf.io/eu7k3>.
- Strydom, Tanya, Giulio V. Dalla Riva, and Timothée Poisot. 2021. “SVD Entropy Reveals the High Complexity of Ecological Networks.” *Frontiers in Ecology and Evolution* 9. <https://doi.org/10.3389/fevo.2021.623141>.
- Trøjelsgaard, Kristian, and Jens M. Olesen. 2016. “Ecological Networks in Motion: Micro- and Macroscopic Variability Across Scales.” *Functional Ecology* 30 (12): 1926–35. <https://doi.org/10.1111/1365-2435.12710>.
- Upham, Nathan S., Jacob A. Esselstyn, and Walter Jetz. 2019. “Inferring the Mammal Tree: Species-Level Sets of Phylogenies for Questions in Ecology, Evolution, and Conservation.” *PLOS Biology* 17 (12): e3000494. <https://doi.org/10.1371/journal.pbio.3000494>.
- Youden, W. J. 1950. “Index for Rating Diagnostic Tests.” *Cancer* 3 (1): 32–35. [https://doi.org/10.1002/1097-0142\(1950\)3:1%3C32::AID-CNCR2820030106%3E3.0.CO;2-3](https://doi.org/10.1002/1097-0142(1950)3:1%3C32::AID-CNCR2820030106%3E3.0.CO;2-3).
- Young, Stephen J., and Edward R. Scheinerman. 2007. “Random Dot Product Graph Models for Social Networks.” In *Algorithms and Models for the Web-Graph*, edited by Anthony Bonato and Fan R. K. Chung, 138–49. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer. [https://doi.org/10.1007/978-3-540-77004-6\\_11](https://doi.org/10.1007/978-3-540-77004-6_11).
- Zhu, Mu, and Ali Ghodsi. 2006. “Automatic Dimensionality Selection from the Scree Plot via the Use of Profile Likelihood.” *Computational Statistics & Data Analysis* 51 (2): 918–30. <https://doi.org/10.1016/j.csda.2005.09.010>.