

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

Tanya Strydom<sup>1,2,‡</sup> Salomé Bouskila<sup>1,‡</sup> Francis Banville<sup>1,3,2</sup> Ceres Barros<sup>4</sup> Dominique Caron<sup>5,2</sup>  
Maxwell J Farrell<sup>6</sup> Marie-Josée Fortin<sup>6</sup> Victoria Hemming<sup>7</sup> Benjamin Mercier<sup>3,2</sup> Laura  
J. Pollock<sup>5,2</sup> Rogini Runghen<sup>8</sup> Giulio V. Dalla Riva<sup>9</sup> Timothée Poisot<sup>1,2</sup>

<sup>1</sup> Département de Sciences Biologiques, Université de Montréal, Montréal, Canada    <sup>2</sup> Quebec Centre for Biodiversity Science, Montréal, Canada    <sup>3</sup> Département de Biologie, Université de Sherbrooke, Sherbrooke, Canada    <sup>4</sup> Department of Forest Resources Management, University of British Columbia, Vancouver, B.C., Canada    <sup>5</sup> Department of Biology, McGill University, Montréal, Canada    <sup>6</sup> Department of Ecology & Evolutionary Biology, University of Toronto, Toronto, Canada    <sup>7</sup> Department of Forest and Conservation Sciences, University of British Columbia, Vancouver, Canada    <sup>8</sup> Centre for Integrative Ecology, School of Biological Sciences, University of Canterbury, Canterbury, New Zealand    <sup>9</sup> School of Mathematics and Statistics, University of Canterbury, Canterbury, New Zealand

<sup>‡</sup> These authors contributed equally to the work

## Correspondance to:

Timothée Poisot — timothee.poisot@umontreal.ca

1. Despite their importance in many ecological processes, collecting data and information on ecological interactions is an exceedingly challenging task. For this reason, large parts of the world have a data deficit when it comes to species interactions, and how the resulting networks are structured. As data collection alone is unlikely to be sufficient, community ecologists must adopt predictive methods.
2. We present a methodological framework that uses graph embedding and transfer learning to assemble a predicted list of trophic interactions of a species pool for which their interactions are unknown. Specifically, we ‘learn’ the information (latent traits) of species from a known interaction network and infer the latent traits of another species pool for which we have no *a priori* interaction data based on their phylogenetic relatedness to species from the known network. The latent traits can then be used to predict interactions and construct an interaction network.
3. Here we assembled a metaweb for Canadian mammals derived from interactions in the European food web, despite only 4% of common species being shared between the two locations. The results of the predictive model are compared against databases of recorded pairwise interactions, showing that we correctly recover 91% of known interactions.
4. The framework itself is robust even when the known network is incomplete or contains spurious interactions making it an ideal candidate as a tool for filling gaps when it comes to species interactions. We provide guidance on how this framework can be adapted by substituting some approaches or predictors in order to make it more generally applicable.

## 1 Introduction

2 There are two core challenges we are faced with in furthering our understanding of ecological networks  
3 across space, particularly at macro-ecologically relevant scales (e.g. Trøjelsgaard & Olesen, 2016). First,  
4 ecological networks within a location are difficult to sample properly (Jordano, 2016a, 2016b), resulting in  
5 a widespread “Eltonian shortfall” (Hortal et al., 2015), *i.e.* a lack of knowledge about inter- and intra-  
6 specific relationships. This first challenge has been, in large part, addressed by the recent emergence of a  
7 suite of methods aiming to predict interactions within *existing* networks, many of which are reviewed in  
8 Strydom, Catchen, et al. (2021). Second, recent analyses based on collected data (Poisot, Bergeron, et al.,  
9 2021) or metadata (Cameron et al., 2019) highlight that ecological networks are currently studied in a  
10 biased subset of space and bioclimates, which impedes our ability to generalize any local understanding of  
11 network structure. Meaning that, although the framework to address incompleteness *within* networks  
12 exists, there would still be regions for which, due to a *lack* of local interaction data, we are unable to infer  
13 potential species interactions.

14 Here, we present a general method to infer potential trophic interactions, relying on the transfer learning  
15 of network representations, specifically by using similarities of species in a biologically/ecologically  
16 relevant proxy space (e.g. shared morphology or ancestry). Transfer learning is a machine learning  
17 methodology that uses the knowledge gained from solving one problem and applying it to a related  
18 (destination) problem (Pan & Yang, 2010; Torrey & Shavlik, 2010). In this instance, we solve the problem  
19 of predicting trophic interactions between species, based on knowledge extracted from another species  
20 pool for which interactions are known by using phylogenetic structure as a medium for transfer. There is a  
21 plurality of measures of species similarities that can be used for inferring *potential* species interactions *i.e.*  
22 metaweb reconstruction (see e.g. Morales-Castilla et al., 2015); however, phylogenetic proximity has  
23 several desirable properties when working at large scales. Gerhold et al. (2015) made the point that  
24 phylogenetic signal captures diversification of characters (large macro-evolutionary process), but not  
25 necessarily community assembly (fine ecological process); Dormann et al. (2010) previously found very  
26 similar conclusions. Interactions tend to reflect a phylogenetic signal because they have a conserved  
27 pattern of evolutionary convergence that encompasses a wide range of ecological and evolutionary  
28 mechanisms (Cavender-Bares et al., 2009; Mouquet et al., 2012), and - most importantly - retain this signal  
29 even if it is obscured at the community scale due to e.g. local conditions (Hutchinson et al., 2017; Poisot &

30 Stouffer, 2018). Finally, species interactions at macro-ecological scales seem to respond mostly to  
31 macro-evolutionary processes (Price, 2003); which is evidenced by the presence of conserved backbones in  
32 food webs (Dalla Riva & Stouffer, 2016; Mora et al., 2018), strong evolutionary signature on prey choice  
33 (Stouffer et al., 2012), and strong phylogenetic signature in food web intervality (Eklöf & Stouffer, 2016).  
34 Phylogenetic reconstruction has also previously been used within the context of ecological networks,  
35 namely understanding ancestral plant-insect interactions (Braga et al., 2021). Taken together, these  
36 considerations suggest that phylogenies can reliably be used to transfer knowledge on species interactions.

37 [Figure 1 about here.]

38 In fig. 1, we provide a methodological overview based on learning the embedding of a metaweb of trophic  
39 interactions for European mammals (known interactions; Maiorano et al., 2020a, 2020b) and, based on  
40 phylogenetic relationships between mammals globally (*i.e.*, phylogenetic tree Upham et al., 2019), infer a  
41 metaweb for the Canadian mammalian species pool (using only a species list *i.e.* interactions are treated as  
42 unknown in this instance). Our case study shows that phylogenetic transfer learning is an effective  
43 approach to the generation of probabilistic metawebs. This showcases that although the components  
44 (species) that make up the Canadian and European communities may be *minimally* shared (the overall  
45 species overlap is less than 4%), if the medium (proxy space) selected in the transfer step is biologically  
46 plausible, we can still effectively learn from the known network and make biologically relevant  
47 predictions of interactions. Indeed, as we detail in the results, when validated against the known (but  
48 fractional) data of trophic interactions present between Canadian mammals, our model achieves a  
49 predictive accuracy of approximately 91%.

## 50 **Method description**

51 The core point of our method is the transfer of knowledge of a known ecological network to predict  
52 interactions between species for another location for which the network is unknown (or partially known)  
53 and is summarized in the grey text boxes in fig. 1. The method we develop is, ecologically speaking, a  
54 “black box,” *i.e.* an algorithm that can be understood mathematically, but whose component parts are not  
55 always directly tied to ecological processes. There is a growing realization in machine learning that  
56 (unintentional) black box algorithms are not necessarily a bad thing (Holm, 2019), as long as their

57 constituent parts can be examined (which is the case with our method). But more importantly, data hold  
58 more information than we might think; as such, even algorithms that are disconnected from a model can  
59 make correct guesses most of the time (Halevy et al., 2009); in fact, in an instance of ecological forecasting  
60 of spatio-temporal systems, model-free approaches (*i.e.* drawing all of their information from the data)  
61 outperformed model-informed ones (Perretti et al., 2013).

## 62 Data used for the case study

63 We use data from the European metaweb assembled by Maiorano et al. (2020a). This was assembled using  
64 data extracted from scientific literature (including published papers, books, and grey literature) from the  
65 last 50 years and includes all terrestrial tetrapods (mammals, breeding birds, reptiles and amphibians)  
66 occurring on the European sub-continent (and Turkey) - with the caveat that only species introduced in  
67 historical times and currently naturalized being included. The European metaweb was filtered using the  
68 Global Biodiversity Information Facility (GBIF) taxonomic backbone (GBIF Secretariat, 2021) so as to  
69 contain only terrestrial and semi-aquatic mammals. As all species had valid matches to the GBIF  
70 taxonomy it was used as the backbone for the remaining reconciliation steps namely, the mammalian  
71 consensus supertree by Upham et al. (2019) (which is used for the knowledge transfer step) and for the  
72 Canadian species list—which was extracted from the International Union for Conservation of Nature  
73 (IUCN) checklist, and corresponds to the same selection criteria that was applied by Maiorano et al.  
74 (2020a) in the European metaweb. After taxonomic cleaning and reconciliation the European metaweb  
75 has 260 species, and the Canadian species pool 163; of these, 17 (about 4% of the total) are shared, and 89  
76 species from Canada (54%) had at least one congeneric species in Europe. The similarity for both species  
77 pools predictably increases with higher taxonomic order, with 19% of shared genera, 47% of shared  
78 families, and 75% of shared orders; for the last point, Canada and Europe each had a single unique order  
79 (*Didelphimorphia* for Canada, *Erinaceomorpha* for Europe).

## 80 Implementation and code availability

81 The entire pipeline is implemented in *Julia* 1.6 (Bezanson et al., 2017) and is available under the  
82 permissive MIT License at <https://osf.io/2zwqm/>. The taxonomic cleanup steps are done using *GBIF.jl*  
83 (Dansereau & Poisot, 2021). The network embedding and analysis is done using *EcologicalNetworks.jl*

84 (Banville et al., 2021; Poisot et al., 2019). The phylogenetic simulations are done using PhyloNetworks.jl  
85 (Solís-Lemus et al., 2017) and Phylo.jl (Reeve et al., 2016). A complete Project.toml file specifying the  
86 full tree of dependencies is available alongside the code. This material also includes a fully annotated copy  
87 of the entire code required to run this project (describing both the intent of the code and discussing some  
88 technical implementation details), a vignette for every step of the process, and a series of Jupyter  
89 notebooks with the text and code. The pipeline can be executed on a laptop in a matter of minutes, and  
90 therefore does not require extensive computational power.

## 91 **Step 1: Learning the origin network representation**

92 The first step in transfer learning is to learn the structure of the original dataset. In order to do so, we rely  
93 on an approach inspired from representational learning, where we learn a *representation* of the metaweb  
94 (in the form of the latent subspaces), rather than a list of interactions (species *a* eats *b*). This approach is  
95 conceptually different from other metaweb-scale predictions (e.g. Albouy et al., 2019), in that the metaweb  
96 representation is easily transferable. Specifically, we use a Random Dot Product Graph model (hereafter  
97 RDPG; S. J. Young & Scheinerman, 2007) to create a number of latent variables that can be combined into  
98 an approximation of the network adjacency matrix. RDPG is known to capture the evolutionary backbone  
99 of food webs (Dalla Riva & Stouffer, 2016), resulting in strong phylogenetic signal in RDPG results; in  
100 other words, the latent variables of an RDPG can be mapped onto a phylogenetic tree, and  
101 phylogenetically similar predators should share phylogenetically similar preys. In addition, recent  
102 advances show that the latent variables produced this way can be used to predict *de novo* interactions.  
103 Interestingly, the latent variables do not need to be produced by decomposing the network itself; in a  
104 recent contribution, Runghen et al. (2021) showed that deep artificial neural networks are able to  
105 reconstruct the left and right subspaces of an RDPG, in order to predict human movement networks from  
106 individual/location metadata and opens up the possibility of using additional metadata as predictors.  
107 The latent variables are created by performing a truncated Singular Value Decomposition (t-SVD; Halko et  
108 al., 2011) on the adjacency matrix. SVD is an appropriate embedding of ecological networks, which has  
109 recently been shown to both capture their complex, emerging properties (Strydom, Dalla Riva, et al., 2021)  
110 and to allow highly accurate prediction of the interactions within a single network (Poisot, Ouellet, et al.,  
111 2021). Under SVD, an adjacency matrix  $\mathbf{A}$  (where  $\mathbf{A}_{m,n} \in \mathbb{B}$  where 1 indicates predation and 0 an absence  
112 thereof) is decomposed into three components resulting in  $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}'$ . Here,  $\Sigma$  is a  $m \times n$  diagonal matrix

113 and contains only singular ( $\sigma$ ) values along its diagonal,  $\mathbf{U}$  is a  $m \times m$  unitary matrix, and  $\mathbf{V}'$  a  $n \times n$   
114 unitary matrix. Truncating the SVD removes additional noise in the dataset by omitting non-zero and/or  
115 smaller  $\sigma$  values from  $\Sigma$  using the rank of the matrix. Under a t-SVD  $\mathbf{A}_{m,n}$  is decomposed so that  $\Sigma$  is a  
116 square  $r \times r$  diagonal matrix (with  $1 \leq r \leq r_{full}$  where  $r_{full}$  is the full rank of  $\mathbf{A}$  and  $r$  the rank at which we  
117 truncate the matrix) containing only non-zero  $\sigma$  values. Additionally,  $\mathbf{U}$  is now an  $m \times r$  semi unitary  
118 matrix and  $\mathbf{V}'$  an  $r \times n$  semi-unitary matrix.

119 The specific rank at which the SVD ought to be truncated is a difficult question. The purpose of SVD is to  
120 remove the noise (expressed at high dimensions) and to focus on the signal (expressed at low dimensions).  
121 In datasets with a clear signal/noise demarcation, a scree plot of  $\Sigma$  can show a sharp drop at the rank where  
122 noise starts (Zhu & Ghodsi, 2006). Because the European metaweb is almost entirely known, the amount  
123 of noise (uncertainty) is low; this is reflected in fig. 2 (left), where the scree plot shows no important drop,  
124 and in fig. 2 (right) where the proportion of variance explained increases smoothly at higher dimensions.  
125 For this reason, we default back to a threshold that explains 60% of the variance in the underlying data,  
126 corresponding to 12 dimensions - i.e. a tradeoff between accuracy and a reduced number of features.

127 An RDPG estimates the probability of observing interactions between nodes (species) as a function of the  
128 nodes' latent variables, and is a way to turn an SVD (which decompose one matrix into three) into two  
129 matrices that can be multiplied to provide an approximation of the network. The latent variables used for  
130 the RDPG, called the left and right subspaces, are defined as  $\mathcal{L} = \mathbf{U}\sqrt{\Sigma}$ , and  $\mathcal{R} = \sqrt{\Sigma}\mathbf{V}'$  – using the full  
131 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 t-SVD  
132 provides a first-order approximation of the network. One advantage of using an RDPG for the network  
133 reconstruction rather than an SVD is that the number of components to estimate decreases; notably, one  
134 does not have to estimate the singular values of the SVD. Furthermore, the two subspaces can be directly  
135 multiplied to yield a network.

136 [Figure 2 about here.]

137 Because RDPG relies on matrix multiplication, the higher dimensions essentially serve to make specific  
138 interactions converge towards 0 or 1; therefore, for reasonably low ranks, there is no guarantee that the  
139 values in the reconstructed network will be within the unit range. In order to determine what constitutes  
140 an appropriate threshold for probability, we performed the RDPG approach on the European metaweb,  
141 and evaluated the probability threshold by treating this as a binary classification problem, specifically

142 assuming that both 0 and 1 in the European metaweb are all true. Given the methodological details given  
143 in Maiorano et al. (2020a) and O'Connor et al. (2020), this seems like a reasonable assumption, although  
144 one that does not hold for all metawebs. We used the thresholding approach presented in Poisot, Ouellet,  
145 et al. (2021), and picked a cutoff that maximized Youden's  $J$  statistic (a measure of the informedness  
146 (trust) of predictions; Youden (1950)); the resulting cutoff was 0.22, and gave an accuracy above 0.99. In  
147 Supp. Mat. 1, we provide several lines of evidence that using the entire network to estimate the threshold  
148 does not lead to overfitting; that using a subset of species would yield the same threshold; that decreasing  
149 the quality of the original data by adding or removing interactions would minimally affect the predictive  
150 accuracy of RDPG applied to the European metaweb; and that the networks reconstructed from artificially  
151 modified data are reconstructed with the correct ecological properties.

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

## 155 **Steps 2 and 3: Transfer learning through phylogenetic relatedness**

156 In order to transfer the knowledge from the European metaweb to the Canadian species pool, we  
157 performed ancestral character estimation using a Brownian motion model, which is a conservative  
158 approach in the absence of strong hypotheses about the nature of phylogenetic signal in the network  
159 decomposition (Litsios & Salamin, 2012). This uses the estimated feature vectors for the European  
160 mammals to create a state reconstruction for all species (conceptually something akin to a trait-based  
161 mammalian phylogeny using latent generality and vulnerability traits) and allows us to impute the  
162 missing (latent) trait data for the Canadian species that are not already in the European network; as we are  
163 focused on predicting contemporary interactions, we only retained the values for the tips of the tree. We  
164 assumed that all traits (*i.e.* the feature vectors for the left and right subspaces) were independent, which is  
165 a reasonable assumption as every trait/dimension added to the t-SVD has an *additive* effect to the one  
166 before it. Note that the Upham et al. (2019) tree itself has some uncertainty associated to inner nodes of  
167 the phylogeny. In this case study we have decided to not propagate this uncertainty as it would complexify  
168 the process. The Brownian motion algorithm returns the *average* value of the trait, and its upper and  
169 lower bounds. Because we do not estimate other parameters of the traits' distributions, we considered that  
170 every species trait is represented as a uniform distribution between these bounds. The choice of the

171 uniform distribution was made because the algorithm returns a minimum and maximum point estimate  
172 for the value, and given this information, the uniform distribution is the one with maximum entropy. Had  
173 all mean parameters estimates been positive, the exponential distribution would have been an alternative,  
174 but this is not the case for the subspaces of an RDPG. In order to examine the consequences of the choice  
175 of distribution, we estimated the variance per latent variable per node to use a Normal distribution; as we  
176 show in Supp. Mat. 2, this decision results in dramatically over-estimating the number and probability of  
177 interactions, and therefore we keep the discussions in the main text to the uniform case. The inferred left  
178 and right subspaces for the Canadian species pool ( $\hat{\mathcal{L}}$  and  $\hat{\mathcal{R}}$ ) have entries that are distributions,  
179 representing the range of values for a given species at a given dimension. These objects represent the  
180 transferred knowledge, which we can use for prediction of the Canadian metaweb.

#### 181 **Step 4: Probabilistic prediction of the destination network**

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

186 [Figure 3 about here.]

187 Specifically, we have adopted the following approach. For every entry in  $\hat{\mathcal{L}}$  and  $\hat{\mathcal{R}}$ , we draw a value from  
188 its distribution. This results in one instance of the possible left ( $\hat{\ell}$ ) and right ( $\hat{r}$ ) subspaces for the  
189 Canadian metaweb. These can be multiplied, to produce one matrix of real values. Because the entries in  
190  $\hat{\ell}$  and  $\hat{r}$  are in the same space where  $\mathcal{L}$  and  $\mathcal{R}$  were originally predicted, it follows that the threshold  $\rho$   
191 estimated for the European metaweb also applies. We use this information to produce one random  
192 Canadian metaweb,  $N = \hat{\mathcal{L}}\hat{\mathcal{R}}' \geq \rho$ . As we can see in (fig. 3), the European and Canadian metawebs are  
193 structurally similar (as would be expected given the biogeographic similarities) and the two (left and right)  
194 subspaces are distinct *i.e.* capturing predation (generality) and prey (vulnerability) latent traits.  
195 Because the intervals around some trait values can be broad (in fact, probably broader than what they  
196 would actually be, see *e.g.* Garland et al., 1999), we repeat the above process  $2 \times 10^5$  times, which results in  
197 a probabilistic metaweb  $P$ , where the probability of an interaction (here conveying our degree of trust that

198 it exists given the inferred trait distributions) is given by the number of times where it appears across all  
199 random draws  $N$ , divided by the number of samples. An interaction with  $P_{i,j} = 1$  means that these two  
200 species were predicted to interact in all  $2 \times 10^5$  random draws.

201 It must be noted that despite bringing in a large amount of information from the European species pool  
202 and interactions, the Canadian metaweb has distinct structural properties. Following an approach similar  
203 to Vermaat et al. (2009), we show in Supp. Mat. 3 that not only can we observe differences in the  
204 multivariate space between the European and Canadian metawebs, we can also observe differences in the  
205 same space between random subgraphs from these networks. These results line up with the studies  
206 spatializing metawebs that have been discussed in the introduction: changes in the species pool are  
207 driving local structural changes in the networks.

## 208 **Data cleanup, discovery, validation, and thresholding**

209 Once the probabilistic metaweb for Canada has been produced, we followed a number of data inflation  
210 steps to finalize it. This step is external to the actual transfer learning framework but rather serves as a  
211 way to augment and validate the predicted metaweb.

212 [Figure 4 about here.]

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

218 Second, we looked for all species in the Canadian pool known to the Global Biotic Interactions (GloBI)  
219 database (Poelen et al., 2014), and extracted their known interactions. Because GloBI aggregates observed  
220 interactions, it is not a *networks* data source, and therefore the only information we can reliably extract  
221 from it is that a species pair *was reported to interact at least once*. This last statement should yet be taken  
222 with caution, as some sources in GloBI (e.g. Thessen & Parr, 2014) are produced through text analysis, and  
223 therefore may not document direct evidence of the interaction. Nevertheless, should the predictive model  
224 work, we would expect that a majority of interactions known to GloBI would also be predicted. We

225 retrieved 366 interactions between mammals from the Canadian species pool from GloBI, 33 of which  
226 were not predicted by the model; this results in a success rate of 91%. After performing this check, we set  
227 the probability of all interactions known to GloBI to 1.

228 Finally, we downloaded the data from Strong & Leroux (2014), who mined various literature sources to  
229 identify trophic interactions in Newfoundland. This dataset documented 25 interactions between  
230 mammals, only two of which were not part of our (Canada-level) predictions, resulting in a success rate of  
231 92%. These two interactions were added to our predicted metaweb with a probability of 1. A comparison  
232 of interaction densities for the inferred metaweb, and the Globi and Newfoundland is shown in fig. 4 and a  
233 table listing all interactions in the predicted Canadian metaweb can be found in the supplementary  
234 material.

235 [Figure 5 about here.]

236 Because the confidence intervals on the inferred trait space are probably over-estimates, we decided to  
237 apply a thresholding step to the interactions after data inflation (see fig. 5 showing the effect of varying the  
238 cutoff on  $P(i \rightarrow j)$ ). Cirtwill & Hambäck (2021) proposed a number of strategies to threshold probabilistic  
239 networks. Their methodology assumes the underlying data to be tag-based sequencing, which represents  
240 interactions as co-occurrences of predator and prey within the same tags; this is conceptually identical to  
241 our Bernoulli-trial based reconstruction of a probabilistic network. We performed a full analysis of the  
242 effect of various cutoffs, and as they either resulted in removing too few interactions, or removing enough  
243 interactions that species started to be disconnected from the network, we set this threshold for a  
244 probability equivalent to 0 to the largest possible value that still allowed all species to have at least one  
245 interaction with a non-zero probability. The need for this slight deviation from the Cirtwill & Hambäck  
246 (2021) methodology highlights the need for additional development on network thresholding.

## 247 Results and discussion

248 [Figure 6 about here.]

249 Using a transfer learning framework we were able to construct a probabilistic metaweb and (as per Dunne,  
250 2006) is a list of potential interactions, meaning that they will not necessarily be realized wherever the two

251 species co-occur. The t-SVD embedding is able to learn relevant ecological features for the network. fig. 6  
252 shows that the first rank correlates linearly with generality and vulnerability (Schoener, 1989), *i.e.* the  
253 number of preys and predators for each species. Importantly, this implies that a rank 1 approximation  
254 represents the configuration model for the metaweb, *i.e.* a set of random networks generated from a given  
255 degree sequence (Park & Newman, 2004). Accounting for the probabilistic nature of the degrees, the rank  
256 1 approximation also represents the *soft* configuration model (van der Hoorn et al., 2018). Both models are  
257 maximum entropy graph models (Garlaschelli et al., 2018), with sharp (all network realizations satisfy the  
258 specified degree sequence) and soft (network realizations satisfy the degree sequence on average) local  
259 constraints, respectively. The (soft) configuration model is an unbiased random graph model widely used  
260 by ecologists in the context of null hypothesis significance testing of network structure (*e.g.* Bascompte et  
261 al., 2003) and can provide informative priors for Bayesian inference of network structure (*e.g.* J.-G. Young  
262 et al., 2021). It is noteworthy that for this metaweb, the relevant information was extracted at the first  
263 rank. Because the first rank corresponds to the leading singular value of the system, the results of fig. 6  
264 have a straightforward interpretation: degree-based processes are the most important in structuring the  
265 mammalian food web.

266 One important aspect in which Europe and Canada differ (despite their comparable bioclimatic  
267 conditions) is the degree of the legacy of human impacts, which have been much longer in Europe.  
268 Nenzén et al. (2014) showed that even at small scales (the Iberian peninsula), mammal food webs retain  
269 the signal of both past climate change and human activity, even when this human activity was orders of  
270 magnitude less important than it is now. Similarly, Yeakel et al. (2014) showed that changes in human  
271 occupation over several centuries can lead to food web collapse. Megafauna in particular seems to be very  
272 sensitive to human arrival (Pires et al., 2015). In short, there is well-substantiated support for the idea that  
273 human footprint affects more than the risk of species extinction (Marco et al., 2018), and can lead to  
274 changes in interaction structure.

275 Cirtwill et al. (2019) showed that network inference techniques based on Bayesian approaches would  
276 perform far better in the presence of an interaction-level informative prior; the desirable properties of such  
277 a prior would be that it is expressed as a probability, preferably representing a Bernoulli event, the value of  
278 which would be representative of relevant biological processes (probability of predation in this case). We  
279 argue that the probability returned at the very last step of our framework may serve as this informative  
280 prior; indeed, the output of our analysis can be used in subsequent steps, also possibly involving expert

281 elicitation to validate some of the most strongly recommended interactions. One important *caveat* to keep  
282 in mind when working with interaction inference is that interactions can never really be true negatives (in  
283 the current state of our methodological framework and data collection limitations); this renders the task of  
284 validating a model through the usual application of binary classification statistics very difficult (although  
285 see Strydom, Catchen, et al., 2021 for a discussion of alternative suggestions). The other way through  
286 which our framework can be improved is by substituting the predictors that are used for transfer. For  
287 example, in the presence of information on species traits that are known to be predictive of species  
288 interactions, one might want to rely on functional rather than phylogenetic distances – in food webs, body  
289 size (and allometrically related variables) has been established as such a variable (Brose et al., 2006); the  
290 identification of relevant functional traits is facilitated by recent methodological developments (Rosado et  
291 al., 2013).

292 Finally, it should be noted that the framework we have presented is amenable to changes lending to  
293 applicability to a broad range of potential scenarios. For example in this case study we have embedded the  
294 original metaweb using t-SVD, because it lends itself to an RDPG reconstruction, which is known to  
295 capture the consequences of evolutionary processes (Dalla Riva & Stouffer, 2016); this being said, there are  
296 other ways to embed graphs (Arsov & Mirceva, 2019; Cai et al., 2017; Cao et al., 2019), which can be used  
297 as alternatives. Regarding the transfer step it is possible to use distinct trees if working with distinct clades  
298 (such as pollination networks) or an alternative measure of similarity (transfer medium) such as  
299 information on foraging (Beckerman et al., 2006), cell-level mechanisms (Boeckaerts et al., 2021), or a  
300 combination of traits and phylogenetic structure (Stock, 2021). Most importantly, although we focus on a  
301 trophic system, it is an established fact that different (non-trophic) interactions do themselves interact with  
302 and influence the outcome of trophic interactions (see e.g. Kawatsu et al., 2021; Kéfi et al., 2012). Future  
303 development of metaweb inference techniques should cover the prediction of multiple interaction types.

304 **Acknowledgements:** We acknowledge that this study was conducted on land within the traditional  
305 unceded territory of the Saint Lawrence Iroquoian, Anishinabewaki, Mohawk, Huron-Wendat, and  
306 Omàmiwininiwak nations. TP, TS, DC, and LP received funding from the Canadian Institute for Ecology  
307 & Evolution. FB is funded by the Institute for Data Valorization (IVADO). TS, SB, and TP are funded by a  
308 donation from the Courtois Foundation. CB was awarded a Mitacs Elevate Fellowship no. IT12391, in  
309 partnership with fRI Research, and also acknowledges funding from Alberta Innovates and the Forest  
310 Resources Improvement Association of Alberta. M-JF acknowledges funding from NSERC Discovery

311 Grant and NSERC CRC. RR is funded by New Zealand's Biological Heritage Ngā Koiora Tuku Iho National  
312 Science Challenge, administered by New Zealand Ministry of Business, Innovation, and Employment. BM  
313 is funded by the NSERC Alexander Graham Bell Canada Graduate Scholarship and the FRQNT master's  
314 scholarship. LP acknowledges funding from NSERC Discovery Grant (NSERC RGPIN-2019-05771). TP  
315 acknowledges financial support from NSERC through the Discovery Grants and Discovery Accelerator  
316 Supplement programs. MJF is supported by an NSERC PDF and an RBC Post-Doctoral Fellowship

317 **Conflict of interest:** The authors have no conflict interests to disclose

318 **Authors' contributions:** TS, SB, and TP designed the study and performed the analysis; GVDR, MF, and  
319 RR provided additional feedback on the analyses. DC, BM, and FB helped with data collection. All  
320 authors contributed to writing and editing the manuscript.

321 **Data availability:** All code and data used in this manuscript is publicly available and archived on OSF  
322 <https://osf.io/2zwqm/> and is currently referenced in the manuscript.

## 323 References

- 324 Albouy, C., Archambault, P., Appeltans, W., Araújo, M. B., Beauchesne, D., Cazelles, K., Cirtwill, A. R.,  
325 Fortin, M.-J., Galiana, N., Leroux, S. J., Pellissier, L., Poisot, T., Stouffer, D. B., Wood, S. A., & Gravel, D.  
326 (2019). The marine fish food web is globally connected. *Nature Ecology & Evolution*, 3(8, 8),  
327 1153–1161. <https://doi.org/10.1038/s41559-019-0950-y>
- 328 Arsov, N., & Mirceva, G. (2019, November 26). *Network Embedding: An Overview*.  
329 <http://arxiv.org/abs/1911.11726>
- 330 Banville, F., Vissault, S., & Poisot, T. (2021). Mangal.jl and EcologicalNetworks.jl: Two complementary  
331 packages for analyzing ecological networks in Julia. *Journal of Open Source Software*, 6(61), 2721.  
332 <https://doi.org/10.21105/joss.02721>
- 333 Bascompte, J., Jordano, P., Melian, C. J., & Olesen, J. M. (2003). The nested assembly of plant-animal  
334 mutualistic networks. *Proceedings of the National Academy of Sciences*, 100(16), 9383–9387.  
335 <https://doi.org/10.1073/pnas.1633576100>
- 336 Beckerman, A. P., Petchey, O. L., & Warren, P. H. (2006). Foraging biology predicts food web complexity.  
337 *Proceedings of the National Academy of Sciences*, 103(37), 13745–13749.

- 338 <https://doi.org/10.1073/pnas.0603039103>
- 339 Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. (2017). Julia: A Fresh Approach to Numerical  
340 Computing. *SIAM Review*, 59(1), 65–98. <https://doi.org/10.1137/141000671>
- 341 Boeckaerts, D., Stock, M., Criel, B., Gerstmans, H., De Baets, B., & Briers, Y. (2021). Predicting  
342 bacteriophage hosts based on sequences of annotated receptor-binding proteins. *Scientific Reports*,  
343 11(1, 1), 1467. <https://doi.org/10.1038/s41598-021-81063-4>
- 344 Braga, M. P., Janz, N., Nylin, S., Ronquist, F., & Landis, M. J. (2021). Phylogenetic reconstruction of  
345 ancestral ecological networks through time for pierid butterflies and their host plants. *Ecology Letters*,  
346 n/a(n/a). <https://doi.org/10.1111/ele.13842>
- 347 Brose, U., Jonsson, T., Berlow, E. L., Warren, P., Banasek-Richter, C., Bersier, L.-F., Blanchard, J. L., Brey,  
348 T., Carpenter, S. R., Blandenier, M.-F. C., Cushing, L., Dawah, H. A., Dell, T., Edwards, F.,  
349 Harper-Smith, S., Jacob, U., Ledger, M. E., Martinez, N. D., Memmott, J., ... Cohen, J. E. (2006).  
350 ConsumerResource Body-Size Relationships in Natural Food Webs. *Ecology*, 87(10), 2411–2417.  
351 [https://doi.org/10.1890/0012-9658\(2006\)87%5B2411:CBRINF%5D2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87%5B2411:CBRINF%5D2.0.CO;2)
- 352 Cai, H., Zheng, V. W., & Chang, K. C.-C. (2017). *A Comprehensive Survey of Graph Embedding: Problems,*  
353 *Techniques and Applications*. <http://arxiv.org/abs/1709.07604>
- 354 Cameron, E. K., Sundqvist, M. K., Keith, S. A., CaraDonna, P. J., Mousing, E. A., Nilsson, K. A., Metcalfe,  
355 D. B., & Classen, A. T. (2019). Uneven global distribution of food web studies under climate change.  
356 *Ecosphere*, 10(3), e02645. <https://doi.org/10.1002/ecs2.2645>
- 357 Cao, R.-M., Liu, S.-Y., & Xu, X.-K. (2019). Network embedding for link prediction: The pitfall and  
358 improvement. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(10), 103102.  
359 <https://doi.org/10.1063/1.5120724>
- 360 Cavender-Bares, J., Kozak, K. H., Fine, P. V. A., & Kembel, S. W. (2009). The merging of community  
361 ecology and phylogenetic biology. *Ecology Letters*, 12(7), 693–715.  
362 <https://doi.org/10.1111/j.1461-0248.2009.01314.x>
- 363 Cirtwill, A. R., Ekl, A., Roslin, T., Wootton, K., & Gravel, D. (2019). A quantitative framework for  
364 investigating the reliability of empirical network construction. *Methods in Ecology and Evolution*, 0.  
365 <https://doi.org/10.1111/2041-210X.13180>

- 366 Cirtwill, A. R., & Hambäck, P. (2021). Building food networks from molecular data: Bayesian or  
367 fixed-number thresholds for including links. *Basic and Applied Ecology*, 50, 67–76.  
368 <https://doi.org/10.1016/j.baae.2020.11.007>
- 369 Dalla Riva, G. V., & Stouffer, D. B. (2016). Exploring the evolutionary signature of food webs' backbones  
370 using functional traits. *Oikos*, 125(4), 446–456. <https://doi.org/10.1111/oik.02305>
- 371 Dansereau, G., & Poisot, T. (2021). SimpleSDMLayers.jl and GBIF.jl: A Framework for Species  
372 Distribution Modeling in Julia. *Journal of Open Source Software*, 6(57), 2872.  
373 <https://doi.org/10.21105/joss.02872>
- 374 Dormann, C. F., Gruber, B., Winter, M., & Herrmann, D. (2010). Evolution of climate niches in European  
375 mammals? *Biology Letters*, 6(2), 229–232. <https://doi.org/10.1098/rsbl.2009.0688>
- 376 Dunne, J. A. (2006). The Network Structure of Food Webs. In J. A. Dunne & M. Pascual (Eds.), *Ecological  
377 networks: Linking structure and dynamics* (pp. 27–86). Oxford University Press.
- 378 Eklöf, A., & Stouffer, D. B. (2016). The phylogenetic component of food web structure and intervality.  
379 *Theoretical Ecology*, 9(1), 107–115. <https://doi.org/10.1007/s12080-015-0273-9>
- 380 Garland, T., JR., Midford, P. E., & Ives, A. R. (1999). An Introduction to Phylogenetically Based Statistical  
381 Methods, with a New Method for Confidence Intervals on Ancestral Values1. *American Zoologist*,  
382 39(2), 374–388. <https://doi.org/10.1093/icb/39.2.374>
- 383 Garlaschelli, D., Hollander, F. den, & Roccaverde, A. (2018). Covariance structure behind breaking of  
384 ensemble equivalence in random graphs. *Journal of Statistical Physics*, 173(3-4), 644–662.  
385 <https://doi.org/10.1007/s10955-018-2114-x>
- 386 GBIF Secretariat. (2021). *GBIF Backbone Taxonomy*. <https://doi.org/10.15468/39omei>
- 387 Gerhold, P., Cahill, J. F., Winter, M., Bartish, I. V., & Prinzing, A. (2015). Phylogenetic patterns are not  
388 proxies of community assembly mechanisms (they are far better). *Functional Ecology*, 29(5), 600–614.  
389 <https://doi.org/10.1111/1365-2435.12425>
- 390 Halevy, A., Norvig, P., & Pereira, F. (2009). The Unreasonable Effectiveness of Data. *IEEE Intelligent  
391 Systems*, 24(2), 8–12. <https://doi.org/10.1109/MIS.2009.36>
- 392 Halko, N., Martinsson, P. G., & Tropp, J. A. (2011). Finding Structure with Randomness: Probabilistic  
393 Algorithms for Constructing Approximate Matrix Decompositions. *SIAM Review*, 53(2), 217–288.

- 394 <https://doi.org/10.1137/090771806>
- 395 Holm, E. A. (2019). In defense of the black box. *Science*, 364(6435), 26–27.
- 396 <https://doi.org/10.1126/science.aax0162>
- 397 Hortal, J., de Bello, F., Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M., & Ladle, R. J. (2015). Seven  
398 Shortfalls that Beset Large-Scale Knowledge of Biodiversity. *Annual Review of Ecology, Evolution, and*  
399 *Systematics*, 46(1), 523–549. <https://doi.org/10.1146/annurev-ecolsys-112414-054400>
- 400 Hutchinson, M. C., Cagua, E. F., & Stouffer, D. B. (2017). Cophylogenetic signal is detectable in pollination  
401 interactions across ecological scales. *Ecology*, n/a–n/a. <https://doi.org/10.1002/ecy.1955>
- 402 Jordano, P. (2016a). Chasing Ecological Interactions. *PLOS Biol*, 14(9), e1002559.  
403 <https://doi.org/10.1371/journal.pbio.1002559>
- 404 Jordano, P. (2016b). Sampling networks of ecological interactions. *Functional Ecology*, 30(12), 1883–1893.  
405 <https://doi.org/10.1111/1365-2435.12763>
- 406 Kawatsu, K., Ushio, M., van Veen, F. J. F., & Kondoh, M. (2021). Are networks of trophic interactions  
407 sufficient for understanding the dynamics of multi-trophic communities? Analysis of a tri-trophic  
408 insect food-web time-series. *Ecology Letters*, 24(3), 543–552. <https://doi.org/10.1111/ele.13672>
- 409 Kéfi, S., Berlow, E. L., Wieters, E. A., Navarrete, S. A., Petchey, O. L., Wood, S. A., Boit, A., Joppa, L. N.,  
410 Lafferty, K. D., Williams, R. J., Martinez, N. D., Menge, B. A., Blanchette, C. A., Iles, A. C., & Brose, U.  
411 (2012). More than a meal... integrating non-feeding interactions into food webs: More than a meal ....  
412 *Ecology Letters*, 15(4), 291–300. <https://doi.org/10.1111/j.1461-0248.2011.01732.x>
- 413 Litsios, G., & Salamin, N. (2012). Effects of Phylogenetic Signal on Ancestral State Reconstruction.  
414 *Systematic Biology*, 61(3), 533–538. <https://doi.org/10.1093/sysbio/syr124>
- 415 Maiorano, L., Montemaggioli, A., Ficetola, G. F., O'Connor, L., & Thuiller, W. (2020a). TETRA-EU 1.0: A  
416 species-level trophic metaweb of European tetrapods. *Global Ecology and Biogeography*, 29(9),  
417 1452–1457. <https://doi.org/10.1111/geb.13138>
- 418 Maiorano, L., Montemaggioli, A., Ficetola, G. F., O'Connor, L., & Thuiller, W. (2020b). *Data from:  
419 Tetra-EU 1.0: A species-level trophic meta-web of European tetrapods* (Version 3, pp. 16596876 bytes)  
420 [Data set]. Dryad. <https://doi.org/10.5061/DRYAD.JM63XSJ7B>
- 421 Marco, M. D., Venter, O., Possingham, H. P., & Watson, J. E. M. (2018). Changes in human footprint drive

- 422 changes in species extinction risk. *Nature Communications*, 9(1), 4621.
- 423 <https://doi.org/10.1038/s41467-018-07049-5>
- 424 Mora, B. B., Gravel, D., Gilarranz, L. J., Poisot, T., & Stouffer, D. B. (2018). Identifying a common backbone  
425 of interactions underlying food webs from different ecosystems. *Nature Communications*, 9(1), 2603.  
426 <https://doi.org/10.1038/s41467-018-05056-0>
- 427 Morales-Castilla, I., Matias, M. G., Gravel, D., & Araújo, M. B. (2015). Inferring biotic interactions from  
428 proxies. *Trends in Ecology & Evolution*, 30(6), 347–356.  
429 <https://doi.org/10.1016/j.tree.2015.03.014>
- 430 Mouquet, N., Devictor, V., Meynard, C. N., Munoz, F., Bersier, L.-F., Chave, J., Couteron, P., Dalecky, A.,  
431 Fontaine, C., Gravel, D., Hardy, O. J., Jabot, F., Lavergne, S., Leibold, M., Mouillot, D., Münkemüller,  
432 T., Pavoine, S., Prinzing, A., Rodrigues, A. S. L., ... Thuiller, W. (2012). Ecophylogenetics: Advances  
433 and perspectives. *Biological Reviews*, 87(4), 769–785.  
434 <https://doi.org/10.1111/j.1469-185X.2012.00224.x>
- 435 Nenzén, H. K., Montoya, D., & Varela, S. (2014). The Impact of 850,000 Years of Climate Changes on the  
436 Structure and Dynamics of Mammal Food Webs. *PLOS ONE*, 9(9), e106651.  
437 <https://doi.org/10.1371/journal.pone.0106651>
- 438 O'Connor, L. M. J., Pollock, L. J., Braga, J., Ficetola, G. F., Maiorano, L., Martinez-Almoyna, C.,  
439 Montemaggioli, A., Ohlmann, M., & Thuiller, W. (2020). Unveiling the food webs of tetrapods across  
440 Europe through the prism of the Eltonian niche. *Journal of Biogeography*, 47(1), 181–192.  
441 <https://doi.org/10.1111/jbi.13773>
- 442 Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data  
443 Engineering*, 22(10), 1345–1359. <https://doi.org/10.1109/TKDE.2009.191>
- 444 Park, J., & Newman, M. E. J. (2004). Statistical mechanics of networks. *Physical Review E*, 70(6), 066117.  
445 <https://doi.org/10.1103/PhysRevE.70.066117>
- 446 Perretti, C. T., Munch, S. B., & Sugihara, G. (2013). Model-free forecasting outperforms the correct  
447 mechanistic model for simulated and experimental data. *Proceedings of the National Academy of  
448 Sciences*, 110(13), 5253–5257. <https://doi.org/10.1073/pnas.1216076110>
- 449 Pires, M. M., Koch, P. L., Fariña, R. A., de Aguiar, M. A. M., dos Reis, S. F., & Guimarães, P. R. (2015).

- 450 Pleistocene megafaunal interaction networks became more vulnerable after human arrival.
- 451 *Proceedings of the Royal Society B: Biological Sciences*, 282(1814), 20151367.
- 452 <https://doi.org/10.1098/rspb.2015.1367>
- 453 Poelen, J. H., Simons, J. D., & Mungall, C. J. (2014). Global biotic interactions: An open infrastructure to  
454 share and analyze species-interaction datasets. *Ecological Informatics*, 24, 148–159.
- 455 <https://doi.org/10.1016/j.ecoinf.2014.08.005>
- 456 Poisot, T., Belisle, Z., Hoebeke, L., Stock, M., & Szefer, P. (2019). EcologicalNetworks.jl - analysing  
457 ecological networks. *Ecography*. <https://doi.org/10.1111/ecog.04310>
- 458 Poisot, T., Bergeron, G., Cazelles, K., Dallas, T., Gravel, D., MacDonald, A., Mercier, B., Violet, C., &  
459 Vissault, S. (2021). Global knowledge gaps in species interaction networks data. *Journal of  
460 Biogeography*, n/a(n/a). <https://doi.org/10.1111/jbi.14127>
- 461 Poisot, T., Cirtwill, A. R., Cazelles, K., Gravel, D., Fortin, M.-J., & Stouffer, D. B. (2016). The structure of  
462 probabilistic networks. *Methods in Ecology and Evolution*, 7(3), 303–312.  
463 <https://doi.org/10.1111/2041-210X.12468>
- 464 Poisot, T., Ouellet, M.-A., Mollentze, N., Farrell, M. J., Becker, D. J., Albery, G. F., Gibb, R. J., Seifert, S. N.,  
465 & Carlson, C. J. (2021, May 31). *Imputing the mammalian virome with linear filtering and singular  
466 value decomposition*. <http://arxiv.org/abs/2105.14973>
- 467 Poisot, T., & Stouffer, D. B. (2018). Interactions retain the co-phylogenetic matching that communities lost.  
468 *Oikos*, 127(2), 230–238. <https://doi.org/10.1111/oik.03788>
- 469 Price, P. W. (2003). *Macroevolutionary theory on macroecological patterns*. Cambridge University Press.
- 470 Reeve, R., Leinster, T., Cobbold, C. A., Thompson, J., Brummitt, N., Mitchell, S. N., & Matthews, L. (2016,  
471 December 8). *How to partition diversity*. <http://arxiv.org/abs/1404.6520>
- 472 Rosado, B. H. P., Dias, A., & de Mattos, E. (2013). Going Back to Basics: Importance of Ecophysiology  
473 when Choosing Functional Traits for Studying Communities and Ecosystems. *Natureza &  
474 Conservaç~ao Revista Brasileira de Conservaç~ao Da Natureza*, 11, 15–22.  
475 <https://doi.org/10.4322/natcon.2013.002>
- 476 Runghen, R., Stouffer, D. B., & Dalla Riva, G. V. (2021). *Exploiting node metadata to predict interactions in  
477 large networks using graph embedding and neural networks*.

- 478        <https://doi.org/10.1101/2021.06.10.447991>
- 479 Schoener, T. W. (1989). Food webs from the small to the large. *Ecology*, 70(6), 1559–1589.
- 480 Solís-Lemus, C., Bastide, P., & Ané, C. (2017). PhyloNetworks: A Package for Phylogenetic Networks.  
481        *Molecular Biology and Evolution*, 34(12), 3292–3298. <https://doi.org/10.1093/molbev/msx235>
- 482 Stock, M. (2021). Pairwise learning for predicting pollination interactions based on traits and phylogeny.  
483        *Ecological Modelling*, 14.
- 484 Stouffer, D. B., Sales-Pardo, M., Sirer, M. I., & Bascompte, J. (2012). Evolutionary Conservation of Species'  
485        Roles in Food Webs. *Science*, 335(6075), 1489–1492. <https://doi.org/10.1126/science.1216556>
- 486 Strong, J. S., & Leroux, S. J. (2014). Impact of Non-Native Terrestrial Mammals on the Structure of the  
487        Terrestrial Mammal Food Web of Newfoundland, Canada. *PLOS ONE*, 9(8), e106264.  
488        <https://doi.org/10.1371/journal.pone.0106264>
- 489 Strydom, T., Catchen, M. D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., Forero-Muñoz,  
490        N. R., Higino, G., Mercier, B., Gonzalez, A., Gravel, D., Pollock, L., & Poisot, T. (2021). A roadmap  
491        towards predicting species interaction networks (across space and time). *Philosophical Transactions of  
492        the Royal Society B: Biological Sciences*, 376(1837), 20210063.  
493        <https://doi.org/10.1098/rstb.2021.0063>
- 494 Strydom, T., Dalla Riva, G. V., & Poisot, T. (2021). SVD Entropy Reveals the High Complexity of Ecological  
495        Networks. *Frontiers in Ecology and Evolution*, 9. <https://doi.org/10.3389/fevo.2021.623141>
- 496 Thessen, A. E., & Parr, C. S. (2014). Knowledge extraction and semantic annotation of text from the  
497        encyclopedia of life. *PloS One*, 9(3), e89550.
- 498 Torrey, L., & Shavlik, J. (2010). Transfer learning. In *Handbook of research on machine learning  
499        applications and trends: Algorithms, methods, and techniques* (pp. 242–264). IGI global.
- 500 Trøjelsgaard, K., & Olesen, J. M. (2016). Ecological networks in motion: Micro- and macroscopic  
501        variability across scales. *Functional Ecology*, 30(12), 1926–1935.  
502        <https://doi.org/10.1111/1365-2435.12710>
- 503 Upham, N. S., Esselstyn, J. A., & Jetz, W. (2019). Inferring the mammal tree: Species-level sets of  
504        phylogenies for questions in ecology, evolution, and conservation. *PLOS Biology*, 17(12), e3000494.  
505        <https://doi.org/10.1371/journal.pbio.3000494>

- 506 van der Hoorn, P., Lippner, G., & Krioukov, D. (2018). Sparse Maximum-Entropy Random Graphs with a  
507 Given Power-Law Degree Distribution. *Journal of Statistical Physics*, 173(3-4), 806–844.  
508 <https://doi.org/10.1007/s10955-017-1887-7>
- 509 Vermaat, J. E., Dunne, J. A., & Gilbert, A. J. (2009). Major dimensions in food-web structure properties.  
510 *Ecology*, 90(1), 278–282. <http://www.ncbi.nlm.nih.gov/pubmed/19294932>
- 511 Yeakel, J. D., Pires, M. M., Rudolf, L., Dominy, N. J., Koch, P. L., Guimarães, P. R., & Gross, T. (2014).  
512 Collapse of an ecological network in Ancient Egypt. *PNAS*, 111(40), 14472–14477.  
513 <https://doi.org/10.1073/pnas.1408471111>
- 514 Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, 3(1), 32–35.  
515 [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)
- 516 Young, J.-G., Cantwell, G. T., & Newman, M. E. J. (2021). Bayesian inference of network structure from  
517 unreliable data. *Journal of Complex Networks*, 8(6). <https://doi.org/10.1093/comnet/cnaa046>
- 518 Young, S. J., & Scheinerman, E. R. (2007). Random Dot Product Graph Models for Social Networks. In A.  
519 Bonato & F. R. K. Chung (Eds.), *Algorithms and Models for the Web-Graph* (pp. 138–149). Springer.  
520 [https://doi.org/10.1007/978-3-540-77004-6\\_11](https://doi.org/10.1007/978-3-540-77004-6_11)
- 521 Zhu, M., & Ghodsi, A. (2006). Automatic dimensionality selection from the scree plot via the use of profile  
522 likelihood. *Computational Statistics & Data Analysis*, 51(2), 918–930.  
523 <https://doi.org/10.1016/j.csda.2005.09.010>

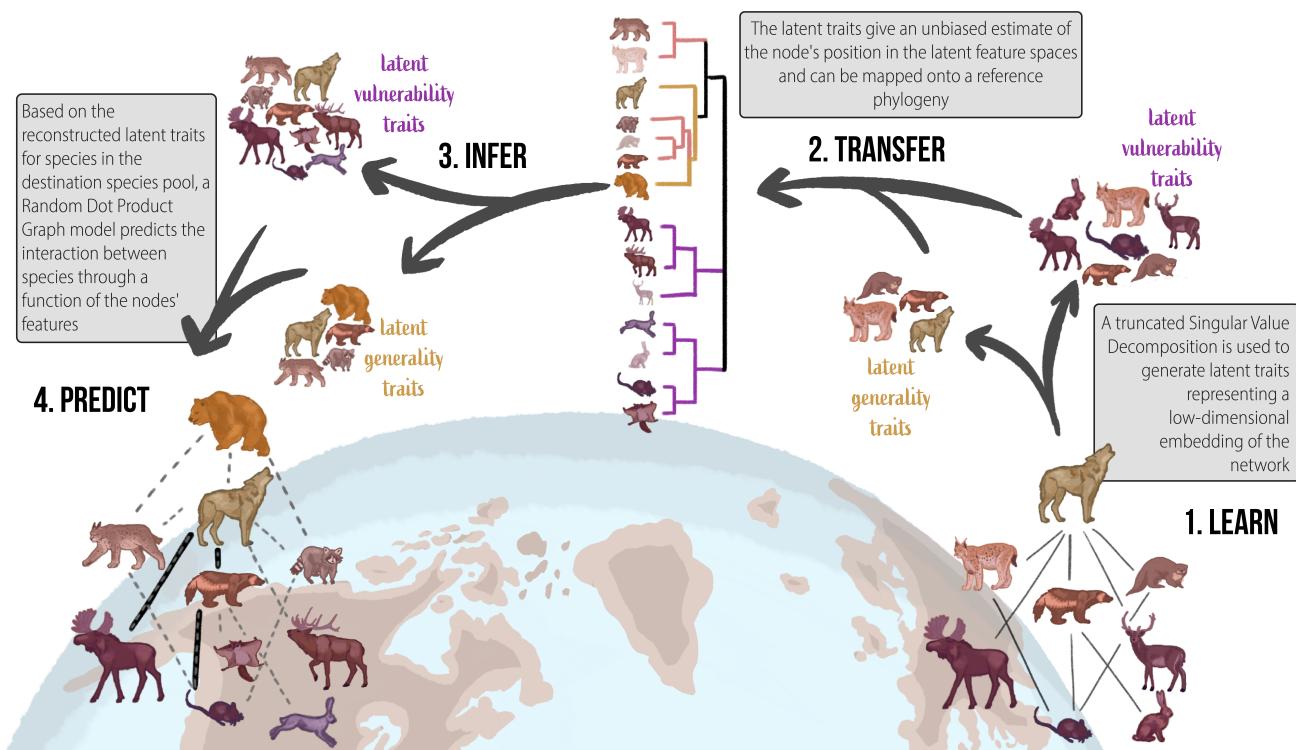


Figure 1: Overview of the phylogenetic transfer learning (and prediction) of species interaction networks. Starting from an initial, known, network, we learn its representation through a graph embedding step (here, a truncated Singular Value Decomposition; Step 1), yielding a series of latent traits (latent vulnerability traits are more representative of species at the lower trophic-level and latent generality traits are more representative of species at higher trophic-levels; *sensu* Schoener (1989)); second, for the destination species pool, we perform ancestral character estimation using a phylogeny (here, using a Brownian model for the latent traits; Step 2); we then sample from the reconstructed distribution of latent traits (Step 3) 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 (Step 4).

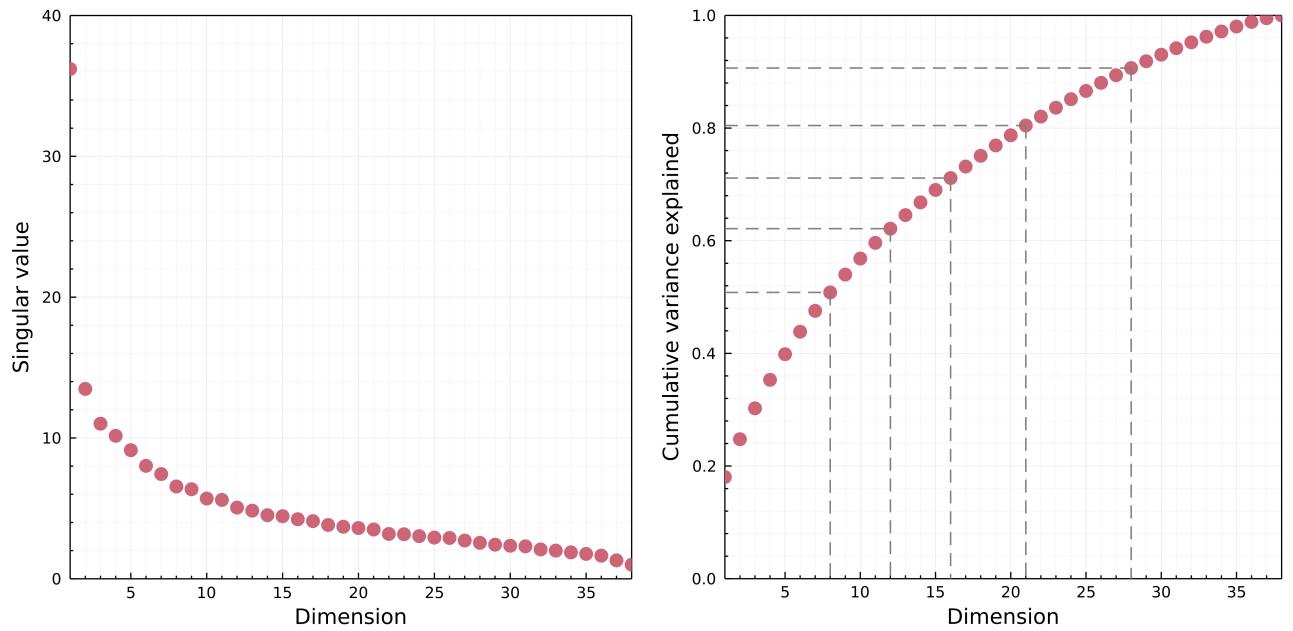


Figure 2: Left: representation of the scree plot of the singular values from the t-SVD on the European metaweb. The scree plot shows no obvious drop in the singular values that may be leveraged to automatically detect a minimal dimension for embedding, after e.g. Zhu & Ghodsi (2006). Right: cumulative fraction of variance explained by each dimension up to the rank of the European metaweb. The grey lines represent cutoffs at 50, 60, ..., 90% 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.

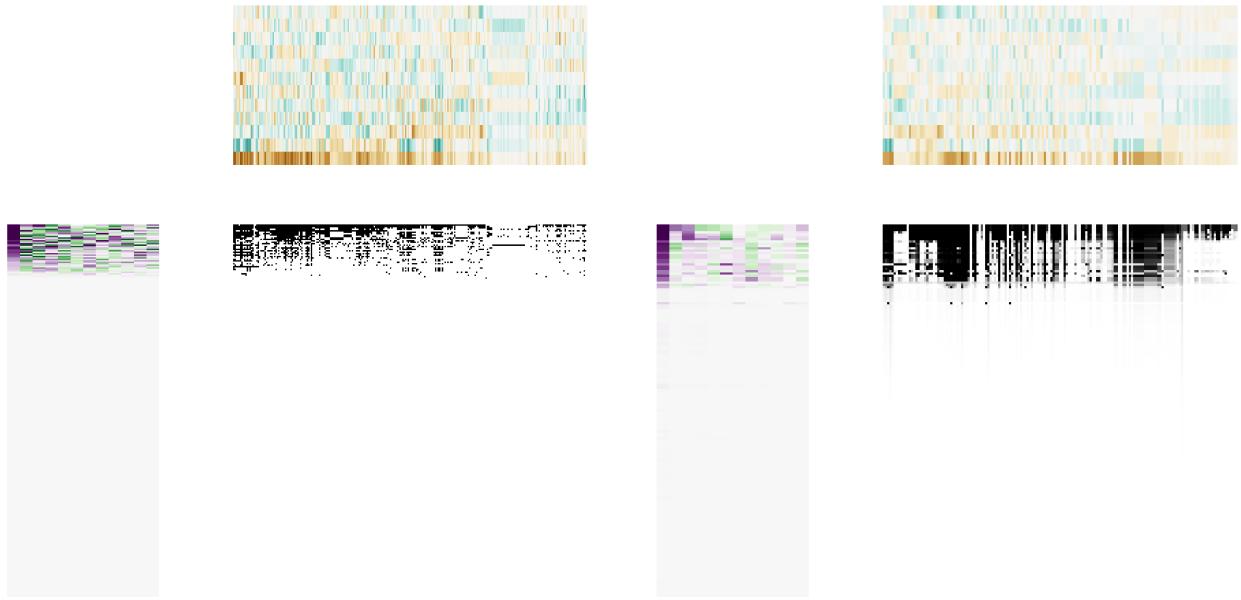


Figure 3: Visual representation of the left (green/purple; left-side matrix) and right (green/brown; top matrix) subspaces, alongside the adjacency matrix of the food web they encode (greyscale). The European metaweb is on the left, and the imputed Canadian metaweb (before data inflation) on the right.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

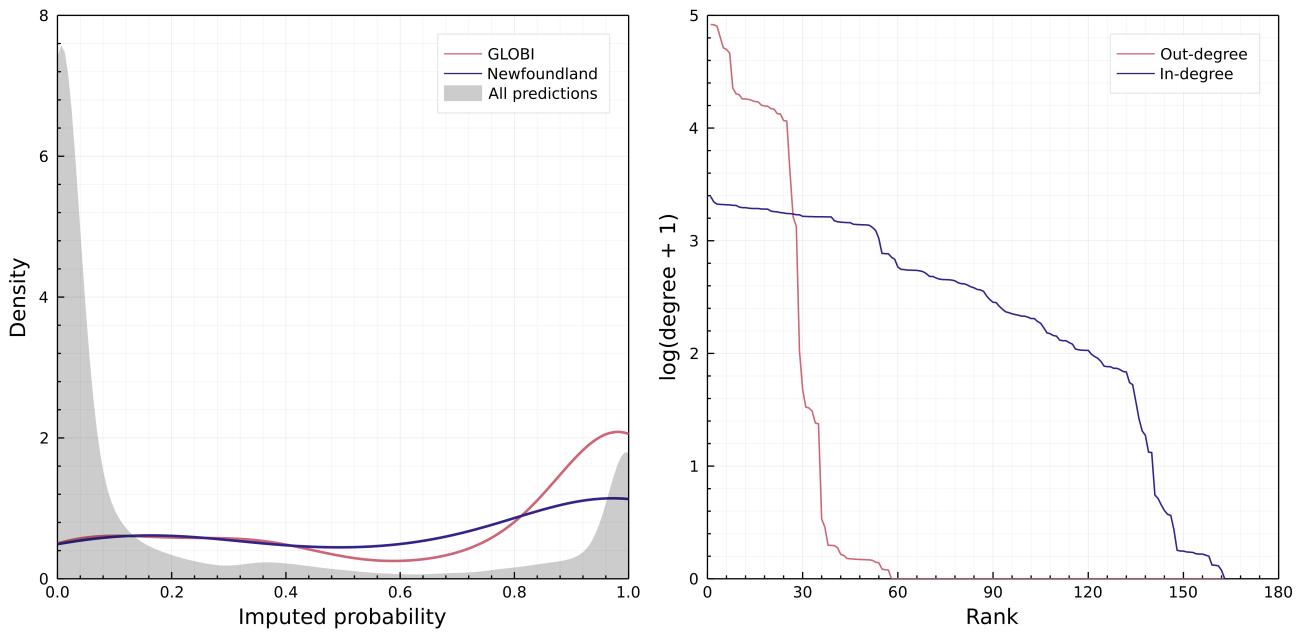


Figure 4: Left: comparison of the probabilities of interactions assigned by the model to all interactions (grey curve), the subset of interactions found in GloBI (red), and in the Strong & Leroux (2014) Newfoundland dataset (blue). The model recovers more interactions with a low probability compared to data mining, which can suggest that collected datasets are biased towards more common or easy to identify interactions. Right: distribution of the in-degree and out-degree of the mammals from Canada in the reconstructed metaweb. This figure describes a flat, relatively short food web, in which there are few predators but a large number of preys.

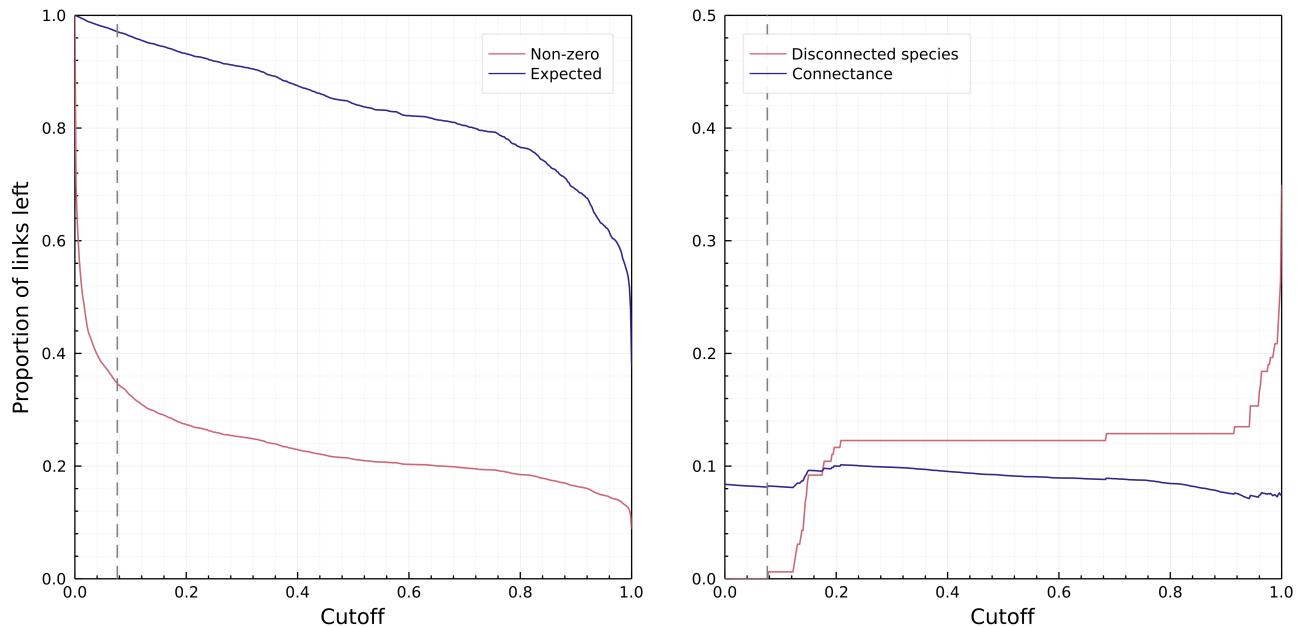


Figure 5: Left: effect of varying the cutoff for probabilities to be considered non-zero on the number of unique links and on  $\hat{L}$ , the probabilistic estimate of the number of links assuming that all interactions are independent. Right: effect of varying the cutoff on the number of disconnected species, and on network connectance. In both panels, the grey line indicates the cutoff  $P(i \rightarrow j) \approx 0.08$  that resulted in the first species losing all of its interactions.

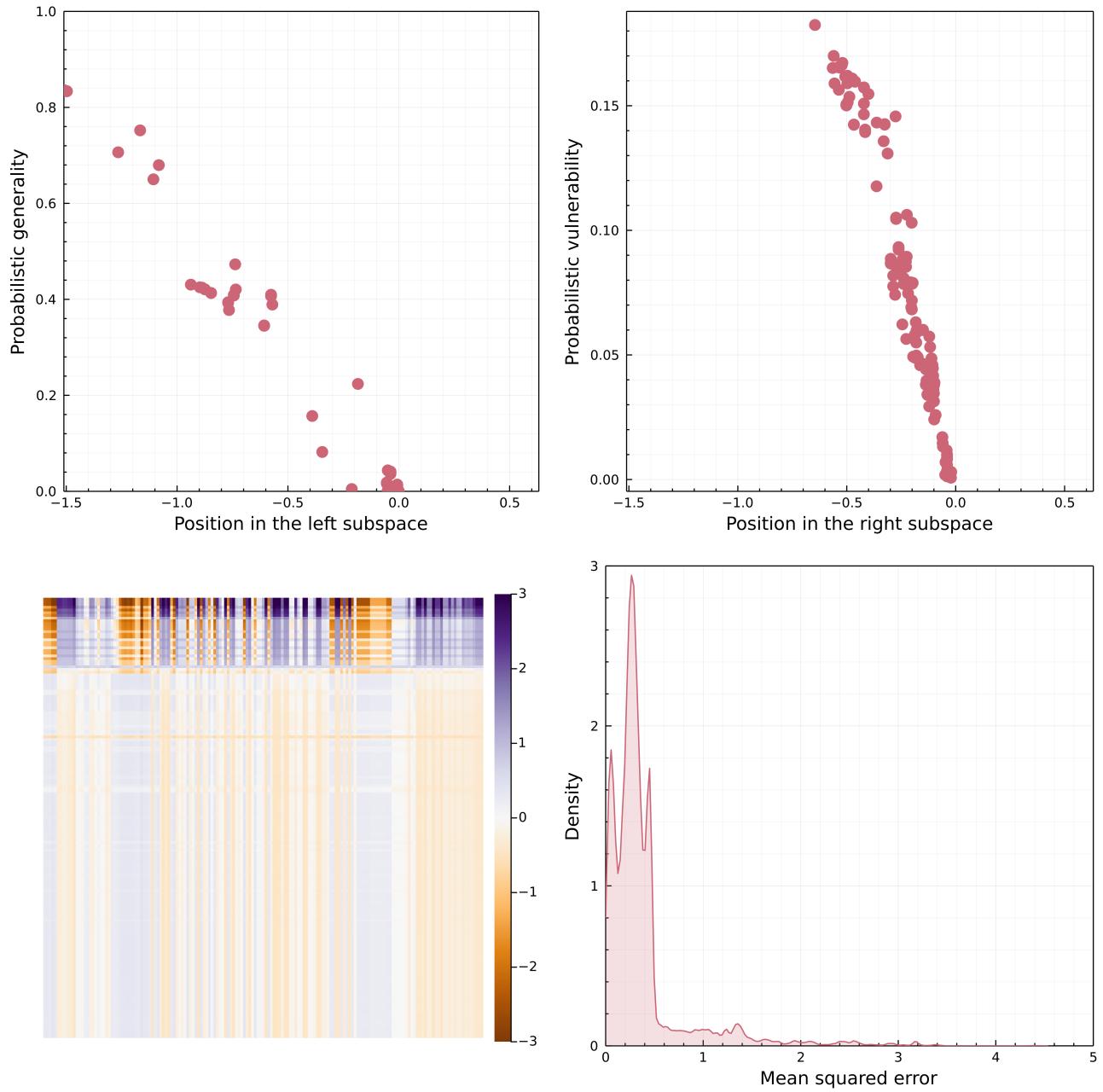


Figure 6: Top: biological significance of the first dimension. Left: there is a linear relationship between the values on the first dimension of the left subspace and the generality, *i.e.* the relative number of preys, *sensu* Schoener (1989). Species with a value of 0 in this subspace are at the bottom-most 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. Bottom: topological consequences of the first dimension. Left: differences in the z-scores of the actual configuration model for the reconstructed network, and the prediction based only on the first dimension. Right: distribution of the differences in the left panel.