

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

Tanya Strydom^{1,2,‡} Salomé Bouskila^{1,‡} Francis Banville^{1,3,2} Ceres Barros⁴ Dominique Caron^{5,2}
Maxwell J Farrell⁶ Marie-Josée Fortin⁶ Victoria Hemming⁷ Benjamin Mercier^{3,2} Laura
J. Pollock^{5,2} Rogini Runghen⁸ Giulio V. Dalla Riva⁹ Timothée Poisot^{1,2}

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

[‡] These authors contributed equally to the work

Correspondance to:

Timothée Poisot — timothee.poisot@umontreal.ca

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. Here, we develop such a method relying on graph embedding and transfer learning to assemble a predicted list of trophic interactions between Canadian mammals. This interaction list is derived from the European food web, despite sharing 4% of common species with Canada. The results of the predictive model are compared against databases of recorded pairwise interactions, showing that we correctly recover 91% of known interactions. We provide guidance on how this method 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, b), resulting in a
5 widespread “Eltonian shortfall” (Hortal *et al.* 2015), *i.e.* a lack of knowledge about inter- and intra- specific
6 relationships. This first challenge has been, in large part, addressed by the recent emergence of a suite of
7 methods aiming to predict interactions within *existing* networks, many of which are reviewed in Strydom
8 *et al.* (2021a). Second, recent analyses based on collected data (Poisot *et al.* 2021a) or metadata (Cameron
9 *et al.* 2019) highlight that ecological networks are currently studied in a biased subset of space and
10 bioclimates, which impedes our ability to generalize any local understanding of network structure.

11 Meaning that, although the framework to address incompleteness *within* networks exists, there would still
12 be regions for which, due to a *lack* of local interaction data, we are unable to infer potential species
13 interactions. Having a general solution for inferring *potential* interactions (despite the unavailability of
14 interaction data) could be the catalyst for significant breakthroughs in our ability to start thinking about
15 species interaction networks over large spatial scales. In a recent overview of the field of ecological
16 network prediction, Strydom *et al.* (2021a) identified two challenges of interest to the prediction of
17 interactions at large scales. First, there is a relative scarcity of relevant data in most places globally –
18 paradoxically, this restricts our ability to infer interactions to locations where inference is perhaps the least
19 required; second, accurate predictions often demand accurate predictors, and the lack of methods that can
20 leverage small amount of data is a serious impediment to our predictive ability globally.

21 Here, we present a general method to recommend potential trophic interactions, relying on the transfer
22 learning of network representations, specifically by using similarities of species in a
23 biologically/ecologically relevant proxy space (e.g. shared morphology or ancestry). Transfer learning is a
24 machine learning methodology that uses the knowledge gained from solving one problem and applying it
25 to a related (destination) problem (Pan & Yang 2010; Torrey & Shavlik 2010). In this instance, we solve the
26 problem of predicting trophic interactions between species, based on knowledge extracted from another
27 species pool for which interactions are known by using phylogenetic structure as a medium for transfer.
28 There is a plurality of measures of species similarities that can be used for inferring *potential* species
29 interactions *i.e.* metaweb reconstruction (see e.g. Morales-Castilla *et al.* 2015); however, phylogenetic

proximity has several desirable properties when working at large scales. Gerhold *et al.* (2015) made the point that phylogenetic signal captures diversification of characters (large macro-evolutionary process), but not necessarily community assembly (fine ecological process); Dormann *et al.* (2010) previously found very similar conclusions. Interactions tend to reflect a phylogenetic signal because they have a conserved pattern of evolutionary convergence that encompasses a wide range of ecological and evolutionary mechanisms (Cavender-Bares *et al.* 2009; Mouquet *et al.* 2012), and - most importantly - retain this signal even when it is not detectable at the community scale (Hutchinson *et al.* 2017; Poisot & Stouffer 2018). Finally, species interactions at macro-ecological scales seem to respond mostly to macro-evolutionary processes (Price 2003); which is evidenced by the presence of conserved backbones in food webs (Dalla Riva & Stouffer 2016; Mora *et al.* 2018), strong evolutionary signature on prey choice (Stouffer *et al.* 2012), and strong phylogenetic signature in food web intervality (Eklöf & Stouffer 2016). Phylogenetic reconstruction has also previously been used within the context of ecological networks, namely understanding ancestral plant-insect interactions (Braga *et al.* 2021). Taken together, these considerations suggest that phylogenies can reliably be used to transfer knowledge on species interactions.

[Figure 1 about here.]

Our methodology is outlined in fig. 1, where we provide an illustration based on learning the embedding of a metaweb of trophic interactions for European mammals (known interactions; Maiorano *et al.* 2020b, a) and, based on phylogenetic relationships between mammals globally (*i.e.*, phylogenetic tree Upham *et al.* 2019), infer a metaweb for the Canadian mammalian species pool (interactions are treated as unknown in this instance). Following the definition of Dunne (2006), a metaweb is a network analogue to the regional species pool; specifically, it is an inventory of all *potential* interactions between species from a spatially delimited area (and so captures the γ diversity of interactions). The metaweb is, therefore, *not* a prediction of the food web at a specific locale within the spatial area it covers, and will have a different structure (notably by having a larger connectance; see *e.g.* Wood *et al.* 2015). These local food webs (which captures the α diversity of interactions) are a subset of the metaweb's species and interactions, and have been called "metaweb realizations" (Poisot *et al.* 2015). Differences between local food web and their metaweb are due to chance, species abundance and co-occurrence, local environmental conditions, and local distribution of functional traits, among others.

Because the metaweb represents the joint effect of functional, phylogenetic, and macroecological

59 processes (Morales-Castilla *et al.* 2015), it holds valuable ecological information. Specifically, it is the
60 “upper bounds” on what the composition of the local networks can be (see e.g. McLeod *et al.* 2021). These
61 local networks, in turn, can be reconstructed given appropriate knowledge of local species composition,
62 providing information on structure of food webs at finer spatial scales. This has been done for example for
63 tree-galler-parasitoid systems (Gravel *et al.* 2018), fish trophic interactions (Albouy *et al.* 2019), tetrapod
64 trophic interactions (O’Connor *et al.* 2020), and crop-pest networks (Grünig *et al.* 2020). Whereas the
65 original metaweb definition, and indeed most past uses of metawebs, was based on the presence/absence
66 of interactions, we focus on *probabilistic* metawebs where interactions are represented as the chance of
67 success of a Bernoulli trial (see e.g. Poisot *et al.* 2016); therefore, not only does our method recommend
68 interactions that may exist, it gives each interaction a score, allowing us to properly weigh them.

69 Our case study shows that phylogenetic transfer learning is an effective approach to the generation of
70 probabilistic metawebs. This showcases that although the components (species) that make up the
71 Canadian and European communities may be *minimally* shared (the overall species overlap is less than
72 4%), if the medium (proxy space) selected in the transfer step is biologically plausible, we can still
73 effectively learn from the known network and make biologically relevant predictions of interactions.

74 Indeed, as we detail in the results, when validated against known but fractional data of trophic
75 interactions between Canadian mammals, our model achieves a predictive accuracy of approximately 91%.
76 It should be reiterated that the framework presented in fig. 1 is amenable to changes; notably, the measure
77 of similarity may not be phylogeny, and can be replaced by information on foraging (Beckerman *et al.*
78 2006), cell-level mechanisms (Boeckaerts *et al.* 2021), or a combination of traits and phylogenetic
79 structure (Stock 2021). Most importantly, although we focus on a trophic system, it is an established fact
80 that different (non-trophic) interactions do themselves interact with and influence the outcome of trophic
81 interactions (Kéfi *et al.* 2012; see e.g. Kawatsu *et al.* 2021). Future development of metaweb inference
82 techniques should cover the prediction of multiple interaction types.

83 Data used for the case study

84 We use data from the European metaweb assembled by Maiorano *et al.* (2020b). This was assembled using
85 data extracted from scientific literature (including published papers, books, and grey literature) from the
86 last 50 years and includes all terrestrial tetrapods (mammals, breeding birds, reptiles and amphibians)

87 occurring on the European sub-continent (and Turkey) - with the caveat that only species introduced in
88 historical times and currently naturalized being included. This metaweb itself is a network of binary (*i.e.*
89 presence/absence), potential two-way interactions between species pairs.

90 We filtered down the European metaweb to create a subgraph corresponding to all mammals by matching
91 species names in the original network to the Global Biodiversity Information Facility (GBIF) taxonomic
92 backbone (GBIF Secretariat 2021) and retaining all those who matched to mammals. This serves a dual
93 purpose 1) to extract only mammals from the European network and 2) to match and standardize species
94 names when aggregating the different data sources further downstream (which is an important
95 consideration when combining datasets (Grenié *et al.* 2021)). All nodes had valid matches to GBIF at this
96 step, and so this backbone is used for all name reconciliation steps as outlined below.

97 The European metaweb represents the knowledge we want to learn and transfer; the phylogenetic
98 similarity of mammals here represents the information for transfer (*i.e.* the transfer medium). We used the
99 mammalian consensus supertree by Upham *et al.* (2019), for which all approximatively 6000 names have
100 been similarly matched to their GBIF valid names. This step allows us to place each node of the
101 mammalian European metaweb in the phylogeny.

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

107 After taxonomic cleaning and reconciliation as outlined in the following sections, the mammalian
108 European metaweb has 260 species, and the Canadian species pool has 163; of these, 17 (about 4% of the
109 total) are shared, and 89 species from Canada (54%) had at least one congeneric species in Europe. The
110 similarity for both species pools predictably increases with higher taxonomic order, with 19% of shared
111 genera, 47% of shared families, and 75% of shared orders; for the last point, Canada and Europe each had a
112 single unique order (*Didelphimorphia* for Canada, *Erinaceomorpha* for Europe).

113 In the following sections, we describe the representational learning step applied to European data, the
114 transfer step through phylogenetic similarity, and the generation of a probabilistic metaweb for the
115 destination species pool.

116 **Method description**

117 The core point of our method is the transfer of knowledge of a known ecological network, in order to
118 predict interactions between species from another location at which the network is unknown (or partially
119 known). In fig. 1, we give a high-level overview of the approach; in the example around which this
120 manuscript is built (leveraging detailed knowledge about binary trophic interactions between Mammalia
121 in Europe to predict the less known trophic interactions between closely phylogenetically related
122 Mammalia in Canada), we use a series of specific steps for network embedding, trait inference, network
123 prediction and thresholding.

124 Specifically, our approach can be summarized as follows: from the known network in Europe, we use a
125 truncated Singular Value Decomposition (t-SVD; Halko *et al.* 2011) to generate latent traits representing a
126 low-dimensional embedding of the network. As an aside, most ecologists are indirectly familiar with SVD:
127 Principal Component Analysis is a special case of SVD, which is more sensitive to numerical instabilities
128 (see notably Shlens 2014). The latent traits give an unbiased estimate of the node's position in the latent
129 feature spaces and can be mapped onto a reference phylogeny (other distance-based measures of species
130 proximity that allow for the inference of features in the latent space can be used, for example the
131 dissimilarity in functional traits). Based on the reconstructed latent traits for species in the destination
132 species pool, a Random Dot Product Graph model (hereafter RDPG; Young & Scheinerman 2007) predicts
133 the interaction between species through a function of the nodes' features through matrix multiplication.
134 Thus, from latent traits and node position, we can infer interactions.

135 The method we develop is, ecologically speaking, a “black box,” *i.e.* an algorithm that can be understood
136 mathematically, but whose component parts are not always directly tied to ecological processes. There is a
137 growing realization in machine learning that (unintentional) black box algorithms are not necessarily a
138 bad thing (Holm 2019), as long as their constituent parts can be examined (which is the case with our
139 method). But more importantly, data hold more information than we might think; as such, even
140 algorithms that are disconnected from the model can make correct guesses most of the time (Halevy *et al.*
141 2009); in fact, in an instance of ecological forecasting of spatio-temporal systems, model-free approaches
142 (*i.e.* drawing all of their information from the data) outperformed model-informed ones (Perretti *et al.*
143 2013).

144 **Implementation and code availability**

145 The entire pipeline is implemented in *Julia* 1.6 (Bezanson *et al.* 2017) and is available under the
146 permissive MIT License at <https://osf.io/2zwqm/>. The taxonomic cleanup steps are done using GBIF.jl
147 (Dansereau & Poisot 2021). The network embedding and analysis is done using EcologicalNetworks.jl
148 (Poisot *et al.* 2019; Banville *et al.* 2021). The phylogenetic simulations are done using PhyloNetworks.jl
149 (Solís-Lemus *et al.* 2017) and Phylo.jl (Reeve *et al.* 2016). A complete Project.toml file specifying the
150 full tree of dependencies is available alongside the code. This material also includes a fully annotated copy
151 of the entire code required to run this project (describing both the intent of the code and discussing some
152 technical implementation details), a vignette for every step of the process, and a series of Jupyter
153 notebooks with the text and code. The pipeline can be executed on a laptop in a matter of minutes, and
154 therefore does not require extensive computational power.

155 **Step 1: Learning the origin network representation**

156 The first step in transfer learning is to learn the structure of the original dataset. In order to do so, we rely
157 on an approach inspired from representational learning, where we learn a *representation* of the metaweb
158 (in the form of the latent subspaces), rather than a list of interactions (species *a* eats *b*). This approach is
159 conceptually different from other metaweb-scale predictions (e.g. Albouy *et al.* 2019), in that the metaweb
160 representation is easily transferable. Specifically, we use RDPG to create a number of latent variables that
161 can be combined into an approximation of the network adjacency matrix. RDPG results are known to
162 have strong phylogenetic signal, and to capture the evolutionary backbone of food webs (Dalla Riva &
163 Stouffer 2016); in other words, the latent variables of an RDPG can be mapped onto a phylogenetic tree,
164 and phylogenetically similar predators should share phylogenetically similar preys. In addition, recent
165 advances show that the latent variables produced this way can be used to predict *de novo* network edges.
166 Interestingly, the latent variables do not need to be produced by decomposing the network itself; in a
167 recent contribution, Runghen *et al.* (2021) showed that deep artificial neural networks are able to
168 reconstruct the left and right subspaces of an RDPG, in order to predict human movement networks from
169 individual/location metadata. This is an exciting opportunity, as it opens up the possibility of using
170 additional metadata as predictors.

171 The latent variables are created by performing a truncated Singular Value Decomposition (t-SVD) on the

172 adjacency matrix. SVD is an appropriate embedding of ecological networks, which has recently been
173 shown to both capture their complex, emerging properties (Strydom *et al.* 2021b) and to allow highly
174 accurate prediction of the interactions within a single network (Poisot *et al.* 2021b). Under SVD, an
175 adjacency matrix \mathbf{A} (where $\mathbf{A}_{m,n} \in \mathbb{B}$ where 1 indicates predation and 0 an absence thereof) is
176 decomposed into three components resulting in $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}'$. Here, Σ is a $m \times n$ diagonal matrix and
177 contains only singular (σ) values along its diagonal, \mathbf{U} is a $m \times m$ unitary matrix, and \mathbf{V}' a $n \times n$ unitary
178 matrix. Truncating the SVD removes additional noise in the dataset by omitting non-zero and/or smaller
179 σ values from Σ using the rank of the matrix. Under a t-SVD $\mathbf{A}_{m,n}$ is decomposed so that Σ is a square $r \times r$
180 diagonal matrix (whith $1 \leq r \leq r_{full}$ where r_{full} is the full rank of \mathbf{A} and r the rank at which we truncate
181 the matrix) containing only non-zero σ values. Additionally, \mathbf{U} is now a $m \times r$ semi unitary matrix and \mathbf{V}'
182 a $n \times r$ semi-unitary matrix.

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

191 An RDPG estimates the probability of observing interactions between nodes (species) as a function of the
192 nodes' latent variables, and is a way to turn a SVD (which decompose one matrix into three) into two
193 matrices that can be multiplied to provide an approximation of the network. The latent variables used for
194 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
195 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
196 provides a first-order approximation of the network. One advantage of using a RDPG rather than a SVD is
197 that the number of components to estimate decreases; notably, one does not have to estimate the singular
198 values of the SVD. Furthermore, the two subspaces can be directly multiplied to yield a network.

199 [Figure 2 about here.]
200 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 for probability, we performed the RDPG approach on the European metaweb, and evaluated the probability 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 details given in Maiorano *et al.* (2020b) and O'Connor *et al.* (2020), this seems like a reasonable assumption, although one that does not hold for all metawebs. We used the thresholding approach presented in Poisot *et al.* (2021b), and picked a cutoff that maximized Youden's *J* statistic (a measure of the informedness (trust) of predictions; Youden (1950)); the resulting cutoff was 0.22, and gave an accuracy above 0.99. In Supp. Mat. 1, we provide several lines of evidence that using the entire network to estimate the threshold does not lead to overfitting; that using a subset of species would yield the same threshold; that decreasing the quality of the original data by adding or removing interactions would minimally affect the predictive accuracy of RDPG applied to the European metaweb; and that the networks reconstructed from artificially modified data are reconstructed with the correct ecological properties.

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

Steps 2 and 3: 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 & Salamin 2012). This uses the estimated feature vectors for the European mammals to create a state reconstruction for all species (conceptually something akin to a trait-based mammalian phylogeny using latent generality and vulnerability traits) and allows us to impute the missing (latent) trait data for the Canadian species that are not already in the European network; as we are focused on predicting contemporary interactions, we only retained the values for the tips of the tree. 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 t-SVD has an *additive* effect to the one before it. Note that the Upham *et al.* (2019) tree itself has some uncertainty associated to inner nodes of

230 the phylogeny. In this case study, we have decided to not propagate this uncertainty, as it would
231 complexify the process. The Brownian motion algorithm returns the *average* value of the trait, and its
232 upper and lower bounds. Because we do not estimate other parameters of the traits' distributions, we
233 considered that every species trait is represented as a uniform distribution between these bounds. The
234 choice of the uniform distribution was made because the algorithm returns a minimum and maximum
235 point estimate for the value, and given this information, the uniform distribution is the one with
236 maximum entropy. Had all mean parameters estimates been positive, the exponential distribution would
237 have been an alternative, but this is not the case for the subspaces of an RDPG. In order to examine the
238 consequences of the choice of distribution, we estimated the variance per latent variable per node to use a
239 Normal distribution; as we show in Supp. Mat. 2, this decision results in dramatically over-estimating the
240 number and probability of interactions, and therefore we keep the discussions in the main text to the
241 uniform case. The inferred left and right subspaces for the Canadian species pool ($\hat{\mathcal{L}}$ and $\hat{\mathcal{R}}$) have entries
242 that are distributions, representing the range of values for a given species at a given dimension.

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

245 **Step 4: Probabilistic prediction of the destination network**

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

250 [Figure 3 about here.]

251 Specifically, we have adopted the following approach. For every entry in $\hat{\mathcal{L}}$ and $\hat{\mathcal{R}}$, we draw a value from
252 its distribution. This results in one instance of the possible left ($\hat{\ell}$) and right (\hat{r}) subspaces for the
253 Canadian metaweb. These can be multiplied, to produce one matrix of real values. Because the entries in
254 $\hat{\ell}$ and \hat{r} are in the same space where \mathcal{L} and \mathcal{R} were originally predicted, it follows that the threshold ρ
255 estimated for the European metaweb also applies. We use this information to produce one random
256 Canadian metaweb, $N = \hat{\mathcal{L}}\hat{\mathcal{R}}' \geq \rho$. As we can see in (fig. 3), the European and Canadian metawebs are

257 structurally similar (as would be expected given the biogeographic similarities) and the two (left and right)
258 subspaces are distinct *i.e.* capturing predation (generality) and prey (vulnerability) latent traits.

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

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

272 **Data cleanup, discovery, validation, and thresholding**

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

276 [Figure 4 about here.]

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

282 Second, we looked for all species in the Canadian pool known to the Global Biotic Interactions (GloBI)
283 database (Poelen *et al.* 2014), 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. Thessen & Parr 2014) are produced through text analysis, and therefore may 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. We retrieved 366 interactions between mammals from the Canadian species pool from GLoBI, 33 of which were not predicted by the model; this results in a success rate of 91%. After performing this check, we set the probability of all interactions known to GLoBI to 1.

Finally, we downloaded the data from Strong & Leroux (2014), who mined various literature sources to identify trophic interactions in Newfoundland. This dataset documented 25 interactions between mammals, only two of which were not part of our (Canada-level) predictions, resulting in a success rate of 92%. These two interactions were added to our predicted metaweb with a probability of 1. A table listing all interactions in the predicted Canadian metaweb can be found in the supplementary material.

[Figure 5 about here.]

Because the confidence intervals on the inferred trait space are probably over-estimates, we decided to apply a thresholding step to the interactions after the data inflation (fig. 5). Cirtwill & Hambäck (2021) proposed 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 too few interactions, or removing enough interactions that species started to be disconnected from the network, we set this threshold for a probability equivalent to 0 to the largest possible value that still allowed all species to have at least one interaction with a non-zero probability. The need for this slight deviation from the Cirtwill & Hambäck (2021) method highlights the need for additional development on network thresholding.

309 **Results and discussion of the case study**

310 In fig. 5, we examine the effect of varying the cutoff on $P(i \rightarrow j)$ on the number of links, species, and
311 connectance. Determining a cutoff using the maximum curvature, or central difference approximation of
312 the second order partial derivative, as suggested by e.g. Cirtwill & Hambäck (2021), results in species being
313 lost, or almost all links being kept. We therefore settled on the value that allowed all species to remain
314 with at least one interaction. This result, in and of itself, suggests that additional methodological
315 developments for the thresholding of probabilistic networks are required.

316 [Figure 6 about here.]

317 The t-SVD embedding is able to learn relevant ecological features for the network. fig. 6 shows that the
318 first rank correlates linearly with generality and vulnerability (Schoener 1989), *i.e.* the number of preys
319 and predators for each species. Importantly, this implies that a rank 1 approximation represents the
320 configuration model for the metaweb, *i.e.* a set of random networks generated from a given degree
321 sequence (Park & Newman 2004). Accounting for the probabilistic nature of the degrees, the rank 1
322 approximation also represents the *soft* configuration model (van der Hoorn *et al.* 2018). Both models are
323 maximum entropy graph models (Garlaschelli *et al.* 2018), with sharp (all network realizations satisfy the
324 specified degree sequence) and soft (network realizations satisfy the degree sequence on average) local
325 constraints, respectively. The (soft) configuration model is an unbiased random graph model widely used
326 by ecologists in the context of null hypothesis significance testing of network structure (e.g. Bascompte *et*
327 *al.* 2003) and can provide informative priors for Bayesian inference of network structure (e.g. Young *et al.*
328 2021). It is noteworthy that for this metaweb, the relevant information was extracted at the first rank.
329 Because the first rank corresponds to the leading singular value of the system, the results of fig. 6 have a
330 straightforward interpretation: degree-based processes are the most important in structuring the
331 mammalian food web.

332 **Discussion**

333 One important aspect in which Europe and Canada differ (despite their comparable bioclimatic
334 conditions) is the degree of the legacy of human impacts, which have been much longer in Europe.

335 Nenzén *et al.* (2014) showed that even at small scales (the Iberian peninsula), mammal food webs retain
336 the signal of both climate change and human activity, even when this human activity was orders of
337 magnitude less important than it is now. Similarly, Yeakel *et al.* (2014) showed that changes in human
338 occupation over several centuries can lead to food web collapse. Megafauna in particular seems to be very
339 sensitive to human arrival (Pires *et al.* 2015). In short, there is well-substantiated support for the idea that
340 human footprint affects more than the risk of species extinction (Marco *et al.* 2018), and can lead to
341 changes in interaction structure. Yet, owing to the inherent plasticity of interactions, there have been
342 documented instances of food webs undergoing rapid collapse/recovery cycles over short periods of time
343 (Pedersen *et al.* 2017). The embedding of a network, in a sense, embeds its macro-evolutionary history,
344 especially as RDPG captures ecological signal (Dalla Riva & Stouffer 2016); at this point, it is important to
345 recall that a metaweb is intended as a catalogue of all potential interactions, which should then be filtered
346 (Morales-Castilla *et al.* 2015). In practice (and in this instance) the reconstructed metaweb will predict
347 interactions that are plausible based on the species' evolutionary history, however some interactions
348 would/would not be realized due to human impact.

349 Dallas *et al.* (2017) suggested that most links in ecological networks may be cryptic, *i.e.* uncommon or
350 otherwise hard to observe. This argument essentially echoes Jordano (2016b): the sampling of ecological
351 interactions is difficult because it requires first the joint observation of two species, and then the
352 observation of their interaction. In addition, it is generally expected that weak or rare links would be more
353 common in networks (Csermely 2004), compared to strong, persistent links; this is notably the case in
354 food chains, wherein many weaker links are key to the stability of a system (Neutel *et al.* 2002). In the
355 light of these observations, the results in fig. 4 are not particularly surprising: we expect to see a surge in
356 these low-probability interactions under a model that has a good predictive accuracy. Because the
357 predictions we generate are by design probabilistic, then one can weigh these rare links appropriately. In a
358 sense, that most ecological interactions are elusive can call for a slightly different approach to sampling:
359 once the common interactions are documented, the effort required in documenting each rare interaction
360 may increase exponentially. Recent proposals suggest that machine learning algorithms, in these
361 situations, can act as data generators (Hoffmann *et al.* 2019): in this perspective, high quality
362 observational data can be supplemented with synthetic data coming from predictive models, which
363 increases the volume of information available for inference. Indeed, Strydom *et al.* (2021a) suggested that
364 knowing the metaweb may render the prediction of local networks easier, because it fixes an “upper

365 bound” on which interactions can exist; indeed, with a probabilistic metaweb, we can consider that the
366 metaweb represents an aggregation of informative priors on the interactions.

367 Related to the last point, Cirtwill *et al.* (2019) showed that network inference techniques based on
368 Bayesian approaches would perform far better in the presence of an interaction-level informative prior;
369 the desirable properties of such a prior would be that it is expressed as a probability, preferably
370 representing a Bernoulli event, the value of which would be representative of relevant biological processes
371 (probability of predation in this case). We argue that the probability returned at the very last step of our
372 framework may serve as this informative prior; indeed, the output of our analysis can be used in
373 subsequent steps, also possibly involving expert elicitation to validate some of the most strongly
374 recommended interactions. One important *caveat* to keep in mind when working with interaction
375 inference is that interactions can never really be true negatives (in the current state of our methodological
376 framework and data collection limitations); this renders the task of validating a model through the usual
377 application of binary classification statistics very difficult (although see Strydom *et al.* 2021a for a
378 discussion of alternative suggestions). The other way through which our framework can be improved is by
379 substituting the predictors that are used for transfer. For example, in the presence of information on
380 species traits that are known to be predictive of species interactions, one might want to rely on functional
381 rather than phylogenetic distances – in food webs, body size (and allometrically related variables) has
382 been established as such a variable (Brose *et al.* 2006); the identification of relevant functional traits is
383 facilitated by recent methodological developments (Rosado *et al.* 2013). It should be noted that Xing &
384 Fayle (2021) highlight phylogenetic relatedness as one of the core components of network comparison at
385 the global scale. In this case study, we have embedded the original metaweb using t-SVD, because it lends
386 itself to an RDPG reconstruction, which is known to capture the consequences of evolutionary processes
387 (Dalla Riva & Stouffer 2016); this being said, there are other ways to embed graphs (Cai *et al.* 2017; Arsov
388 & Mirceva 2019; Cao *et al.* 2019), which can be used as alternatives.

389 As Herbert (1965) rightfully pointed out, “[y]ou can’t draw neat lines around planet-wide problems”; in
390 this regard, our approach (and indeed, any inference of a metaweb at large scales) must contend with
391 several interesting and interwoven families of problems. The first is the limit of the metaweb to embed
392 and transfer. If the initial metaweb is too narrow in scope, notably from a taxonomic point of view, the
393 chances of finding another area with enough related species to make a reliable inference decreases; this
394 would likely be indicated by large confidence intervals during ancestral character estimation, but the lack

of well documented metawebs is currently preventing the development of more concrete guidelines. The question of phylogenetic relatedness and dispersal is notably true if the metaweb is assembled in an area with mostly endemic species, and as with every predictive algorithm, there is room for the application of our best ecological judgement. Conversely, the metaweb should be reliably filled, which assumes that the S^2 interactions in a pool of S species have been examined, either through literature surveys or expert elicitation. Supp. Mat. 1 provides some guidance as to the type of sampling effort that should be prioritized. Although RDPG was able to maintain very high predictive power when interactions were missing, the addition of false positive interactions was immediately detected; this suggests that it may be appropriate to err on the side of “too many” interactions when constructing the initial metaweb to be transferred. The second series of problems are related to determining which area should be used to infer the new metaweb in, as this determines the species pool that must be used. In our application, we focused on the mammals of Canada. The upside of this approach is that information at the country level is likely to be required by policy makers and stakeholders for their biodiversity assessment, as each country tends to set goals at the national level (Buxton *et al.* 2021) for which quantitative instruments are designed (Turak *et al.* 2017), with specific strategies often enacted at smaller scales (Ray *et al.* 2021). And yet, we do not really have a satisfying answer to the question of “where does a food web stop?”; the current most satisfying solutions involve examining the spatial consistency of network area relationships (see e.g. Galiana *et al.* 2018, 2019, 2021; Fortin *et al.* 2021), which is of course impossible in the absence of enough information about the network itself. This suggests that an *a posteriori* refinement of the results may be required, based on a downscaling of the metaweb. The final family of problems relates less to the availability of data or quantitative tools, and more to the praxis of spatial ecology. Operating under the context of national divisions, in large parts of the world, reflects nothing more than the legacy of settler colonialism. Indeed, the use of ecological data is not an apolitical act (Nost & Goldstein 2021), as data infrastructures tend to be designed to answer questions within national boundaries, and their use both draws upon and reinforces territorial statecraft; as per Machen & Nost (2021), this is particularly true when the output of “algorithmic thinking” (e.g. relying on machine learning to generate knowledge) can be re-used for governance (e.g. enacting conservation decisions at the national scale). We therefore recognize that methods such as we propose operate under the framework that contributed to the ongoing biodiversity crisis (Adam 2014), reinforced environmental injustice (Choudry 2013; Domínguez & Luoma 2020), and on Turtle Island especially, should be replaced by Indigenous principles of land management

425 (Eichhorn *et al.* 2019; No'kmaq *et al.* 2021). As we see AI/ML being increasingly mobilized to generate
426 knowledge that is lacking for conservation decisions (*e.g.* Lamba *et al.* 2019; Mosebo Fernandes *et al.*
427 2020), our discussion of these tools need to go beyond the technical, and into the governance
428 consequences they can have.

429 **Acknowledgements:** We acknowledge that this study was conducted on land within the traditional
430 unceded territory of the Saint Lawrence Iroquoian, Anishinabewaki, Mohawk, Huron-Wendat, and
431 Omàmiwininiwak nations. TP, TS, DC, and LP received funding from the Canadian Institute for Ecology &
432 Evolution. FB is funded by the Institute for Data Valorization (IVADO). TS, SB, and TP are funded by a
433 donation from the Courtois Foundation. CB was awarded a Mitacs Elevate Fellowship no. IT12391, in
434 partnership with fRI Research, and also acknowledges funding from Alberta Innovates and the Forest
435 Resources Improvement Association of Alberta. M-JF acknowledges funding from NSERC Discovery
436 Grant and NSERC CRC. RR is funded by New Zealand's Biological Heritage Ngā Koiora Tuku Iho
437 National Science Challenge, administered by New Zealand Ministry of Business, Innovation, and
438 Employment. BM is funded by the NSERC Alexander Graham Bell Canada Graduate Scholarship and the
439 FRQNT master's scholarship. LP acknowledges funding from NSERC Discovery Grant (NSERC
440 RGPIN-2019-05771). TP acknowledges financial support from NSERC through the Discovery Grants and
441 Discovery Accelerator Supplement programs.

442 References

- 443 Adam, R. (2014). *Elephant treaties: The Colonial legacy of the biodiversity crisis*. UPNE.
- 444 Albouy, C., Archambault, P., Appeltans, W., Araújo, M.B., Beauchesne, D., Cazelles, K., *et al.* (2019). The
445 marine fish food web is globally connected. *Nature Ecology & Evolution*, 3, 1153–1161.
- 446 Arsov, N. & Mirceva, G. (2019). *Network Embedding: An Overview*. Available at:
447 <http://arxiv.org/abs/1911.11726>. Last accessed.
- 448 Banville, F., Vissault, S. & Poisot, T. (2021). Mangal.jl and EcologicalNetworks.jl: Two complementary
449 packages for analyzing ecological networks in Julia. *Journal of Open Source Software*, 6, 2721.
- 450 Bascompte, J., Jordano, P., Melian, C.J. & Olesen, J.M. (2003). The nested assembly of plant-animal
451 mutualistic networks. *Proceedings of the National Academy of Sciences*, 100, 9383–9387.

- 452 Beckerman, A.P., Petchey, O.L. & Warren, P.H. (2006). Foraging biology predicts food web complexity.
- 453 *Proceedings of the National Academy of Sciences*, 103, 13745–13749.
- 454 Bezanson, J., Edelman, A., Karpinski, S. & Shah, V. (2017). Julia: A Fresh Approach to Numerical
- 455 Computing. *SIAM Review*, 59, 65–98.
- 456 Boeckaerts, D., Stock, M., Criel, B., Gerstmans, H., De Baets, B. & Briers, Y. (2021). Predicting
- 457 bacteriophage hosts based on sequences of annotated receptor-binding proteins. *Scientific Reports*, 11,
- 458 1467.
- 459 Braga, M.P., Janz, N., Nylin, S., Ronquist, F. & Landis, M.J. (2021). Phylogenetic reconstruction of ancestral
- 460 ecological networks through time for pierid butterflies and their host plants. *Ecology Letters*, n/a.
- 461 Brose, U., Jonsson, T., Berlow, E.L., Warren, P., Banasek-Richter, C., Bersier, L.-F., *et al.* (2006).
- 462 ConsumerResource Body-Size Relationships in Natural Food Webs. *Ecology*, 87, 2411–2417.
- 463 Buxton, R.T., Bennett, J.R., Reid, A.J., Shulman, C., Cooke, S.J., Francis, C.M., *et al.* (2021). Key
- 464 information needs to move from knowledge to action for biodiversity conservation in Canada.
- 465 *Biological Conservation*, 256, 108983.
- 466 Cai, H., Zheng, V.W. & Chang, K.C.-C. (2017). *A Comprehensive Survey of Graph Embedding: Problems,*
- 467 *Techniques and Applications*. Available at: <http://arxiv.org/abs/1709.07604>. Last accessed.
- 468 Cameron, E.K., Sundqvist, M.K., Keith, S.A., CaraDonna, P.J., Mousing, E.A., Nilsson, K.A., *et al.* (2019).
- 469 Uneven global distribution of food web studies under climate change. *Ecosphere*, 10, e02645.
- 470 Cao, R.-M., Liu, S.-Y. & Xu, X.-K. (2019). Network embedding for link prediction: The pitfall and
- 471 improvement. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29, 103102.
- 472 Cavender-Bares, J., Kozak, K.H., Fine, P.V.A. & Kembel, S.W. (2009). The merging of community ecology
- 473 and phylogenetic biology. *Ecology Letters*, 12, 693–715.
- 474 Choudry, A. (2013). Saving biodiversity, for whom and for what? Conservation NGOs, complicity,
- 475 colonialism and conquest in an era of capitalist globalization. In: *NGOization: Complicity,*
- 476 *contradictions and prospects*. Bloomsbury Publishing, pp. 24–44.
- 477 Cirtwill, A.R., Ekl, A., Roslin, T., Wootton, K. & Gravel, D. (2019). A quantitative framework for
- 478 investigating the reliability of empirical network construction. *Methods in Ecology and Evolution*, 0.

- 479 Cirtwill, A.R. & Hambäck, P. (2021). Building food networks from molecular data: Bayesian or
480 fixed-number thresholds for including links. *Basic and Applied Ecology*, 50, 67–76.
- 481 Csermely, P. (2004). Strong links are important, but weak links stabilize them. *Trends in Biochemical
482 Sciences*, 29, 331–334.
- 483 Dalla Riva, G.V. & Stouffer, D.B. (2016). Exploring the evolutionary signature of food webs' backbones
484 using functional traits. *Oikos*, 125, 446–456.
- 485 Dallas, T., Park, A.W. & Drake, J.M. (2017). Predicting cryptic links in host-parasite networks. *PLOS
486 Computational Biology*, 13, e1005557.
- 487 Dansereau, G. & Poisot, T. (2021). SimpleSDMLayers.jl and GBIF.jl: A Framework for Species Distribution
488 Modeling in Julia. *Journal of Open Source Software*, 6, 2872.
- 489 Domínguez, L. & Luoma, C. (2020). Decolonising Conservation Policy: How Colonial Land and
490 Conservation Ideologies Persist and Perpetuate Indigenous Injustices at the Expense of the
491 Environment. *Land*, 9, 65.
- 492 Dormann, C.F., Gruber, B., Winter, M. & Herrmann, D. (2010). Evolution of climate niches in European
493 mammals? *Biology Letters*, 6, 229–232.
- 494 Dunne, J.A. (2006). The Network Structure of Food Webs. In: *Ecological networks: Linking structure and
495 dynamics* (eds. Dunne, J.A. & Pascual, M.). Oxford University Press, pp. 27–86.
- 496 Eichhorn, M.P., Baker, K. & Griffiths, M. (2019). Steps towards decolonising biogeography. *Frontiers of
497 Biogeography*, 12, 1–7.
- 498 Eklöf, A. & Stouffer, D.B. (2016). The phylogenetic component of food web structure and intervality.
499 *Theoretical Ecology*, 9, 107–115.
- 500 Fortin, M.-J., Dale, M.R.T. & Brimacombe, C. (2021). Network ecology in dynamic landscapes. *Proceedings
501 of the Royal Society B: Biological Sciences*, 288, rspb.2020.1889, 20201889.
- 502 Galiana, N., Barros, C., Braga, J., Ficetola, G.F., Maiorano, L., Thuiller, W., et al. (2021). The spatial scaling
503 of food web structure across European biogeographical regions. *Ecography*, n/a.
- 504 Galiana, N., Hawkins, B.A. & Montoya, J.M. (2019). The geographical variation of network structure is
505 scale dependent: Understanding the biotic specialization of hostparasitoid networks. *Ecography*, 42,
506 1175–1187.

- 507 Galiana, N., Lurgi, M., Claramunt-López, B., Fortin, M.-J., Leroux, S., Cazelles, K., *et al.* (2018). The
508 spatial scaling of species interaction networks. *Nature Ecology & Evolution*, 2, 782–790.
- 509 Garland, T., JR., Midford, P.E. & Ives, A.R. (1999). An Introduction to Phylogenetically Based Statistical
510 Methods, with a New Method for Confidence Intervals on Ancestral Values1. *American Zoologist*, 39,
511 374–388.
- 512 Garlaschelli, D., Hollander, F. den & Roccaverde, A. (2018). Covariance structure behind breaking of
513 ensemble equivalence in random graphs. *Journal of Statistical Physics*, 173, 644–662.
- 514 GBIF Secretariat. (2021). GBIF Backbone Taxonomy.
- 515 Gerhold, P., Cahill, J.F., Winter, M., Bartish, I.V. & Prinzing, A. (2015). Phylogenetic patterns are not
516 proxies of community assembly mechanisms (they are far better). *Functional Ecology*, 29, 600–614.
- 517 Gravel, D., Baiser, B., Dunne, J.A., Kopelke, J.-P., Martinez, N.D., Nyman, T., *et al.* (2018). Bringing Elton
518 and Grinnell together: A quantitative framework to represent the biogeography of ecological
519 interaction networks. *Ecography*, 0.
- 520 Grenié, M., Berti, E., Carvajal-Quintero, J.D., Winter, M. & Sagouis, A. (2021). Harmonizing taxon names
521 in biodiversity data: A review of tools, databases, and best practices.
- 522 Grünig, M., Mazzi, D., Calanca, P., Karger, D.N. & Pellissier, L. (2020). Crop and forest pest metawebs shift
523 towards increased linkage and suitability overlap under climate change. *Communications Biology*, 3,
524 1–10.
- 525 Halevy, A., Norvig, P. & Pereira, F. (2009). The Unreasonable Effectiveness of Data. *IEEE Intelligent
526 Systems*, 24, 8–12.
- 527 Halko, N., Martinsson, P.G. & Tropp, J.A. (2011). Finding Structure with Randomness: Probabilistic
528 Algorithms for Constructing Approximate Matrix Decompositions. *SIAM Review*, 53, 217–288.
- 529 Herbert, F. (1965). *Dune*. 1st edn. Chilton Book Company, Philadelphia.
- 530 Hoffmann, J., Bar-Sinai, Y., Lee, L.M., Andrejevic, J., Mishra, S., Rubinstein, S.M., *et al.* (2019). Machine
531 learning in a data-limited regime: Augmenting experiments with synthetic data uncovers order in
532 crumpled sheets. *Science Advances*, 5, eaau6792.
- 533 Holm, E.A. (2019). In defense of the black box. *Science*, 364, 26–27.

- 534 Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M. & Ladle, R.J. (2015). Seven
535 Shortfalls that Beset Large-Scale Knowledge of Biodiversity. *Annual Review of Ecology, Evolution, and*
536 *Systematics*, 46, 523–549.
- 537 Hutchinson, M.C., Cagua, E.F. & Stouffer, D.B. (2017). Cophylogenetic signal is detectable in pollination
538 interactions across ecological scales. *Ecology*, n/a–n/a.
- 539 Jordano, P. (2016a). Chasing Ecological Interactions. *PLOS Biol*, 14, e1002559.
- 540 Jordano, P. (2016b). Sampling networks of ecological interactions. *Functional Ecology*, 30, 1883–1893.
- 541 Kawatsu, K., Ushio, M., van Veen, F.J.F. & Kondoh, M. (2021). Are networks of trophic interactions
542 sufficient for understanding the dynamics of multi-trophic communities? Analysis of a tri-trophic
543 insect food-web time-series. *Ecology Letters*, 24, 543–552.
- 544 Kéfi, S., Berlow, E.L., Wieters, E.A., Navarrete, S.A., Petchey, O.L., Wood, S.A., et al. (2012). More than a
545 meal... integrating non-feeding interactions into food webs: More than a meal *Ecology Letters*, 15,
546 291–300.
- 547 Lamba, A., Cassey, P., Segaran, R.R. & Koh, L.P. (2019). Deep learning for environmental conservation.
548 *Current Biology*, 29, R977–R982.
- 549 Litsios, G. & Salamin, N. (2012). Effects of Phylogenetic Signal on Ancestral State Reconstruction.
550 *Systematic Biology*, 61, 533–538.
- 551 Machen, R. & Nost, E. (2021). Thinking algorithmically: The making of hegemonic knowledge in climate
552 governance. *Transactions of the Institute of British Geographers*, 46, 555–569.
- 553 Maiorano, L., Montemaggiore, A., Ficetola, G.F., O'Connor, L. & Thuiller, W. (2020a). Data from: Tetra-EU
554 1.0: A species-level trophic meta-web of European tetrapods.
- 555 Maiorano, L., Montemaggiori, A., Ficetola, G.F., O'Connor, L. & Thuiller, W. (2020b). TETRA-EU 1.0: A
556 species-level trophic metaweb of European tetrapods. *Global Ecology and Biogeography*, 29, 1452–1457.
- 557 Marco, M.D., Venter, O., Possingham, H.P. & Watson, J.E.M. (2018). Changes in human footprint drive
558 changes in species extinction risk. *Nature Communications*, 9, 4621.
- 559 McLeod, A., Leroux, S.J., Gravel, D., Chu, C., Cirtwill, A.R., Fortin, M.-J., et al. (2021). Sampling and
560 asymptotic network properties of spatial multi-trophic networks. *Oikos*, n/a.

- 561 Mora, B.B., Gravel, D., Gilarranz, L.J., Poisot, T. & Stouffer, D.B. (2018). Identifying a common backbone of
562 interactions underlying food webs from different ecosystems. *Nature Communications*, 9, 2603.
- 563 Morales-Castilla, I., Matias, M.G., Gravel, D. & Araújo, M.B. (2015). Inferring biotic interactions from
564 proxies. *Trends in Ecology & Evolution*, 30, 347–356.
- 565 Mosebo Fernandes, A.C., Quintero Gonzalez, R., Lenihan-Clarke, M.A., Leslie Trotter, E.F. & Jokar
566 Arsanjani, J. (2020). Machine Learning for Conservation Planning in a Changing Climate.
567 *Sustainability*, 12, 7657.
- 568 Mouquet, N., Devictor, V., Meynard, C.N., Munoz, F., Bersier, L.-F., Chave, J., et al. (2012).
569 Ecophylogenetics: Advances and perspectives. *Biological Reviews*, 87, 769–785.
- 570 Nenzén, H.K., Montoya, D. & Varela, S. (2014). The Impact of 850,000 Years of Climate Changes on the
571 Structure and Dynamics of Mammal Food Webs. *PLOS ONE*, 9, e106651.
- 572 Neutel, A.-M., Heesterbeek, J.A.P. & de Ruiter, P.C. (2002). Stability in Real Food Webs: Weak Links in
573 Long Loops. *Science*, 296, 1120–1123.
- 574 No'kmaq, M., Marshall, A., Beazley, K.F., Hum, J., joudry, shalan, Papadopoulos, A., et al. (2021).
575 “Awakening the sleeping giant”: Re-Indigenization principles for transforming biodiversity
576 conservation in Canada and beyond. *FACETS*, 6, 839–869.
- 577 Nost, E. & Goldstein, J.E. (2021). A political ecology of data. *Environment and Planning E: Nature and
578 Space*, 25148486211043503.
- 579 O'Connor, L.M.J., Pollock, L.J., Braga, J., Ficetola, G.F., Maiorano, L., Martinez-Almoyna, C., et al. (2020).
580 Unveiling the food webs of tetrapods across Europe through the prism of the Eltonian niche. *Journal of
581 Biogeography*, 47, 181–192.
- 582 Pan, S.J. & Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data
583 Engineering*, 22, 1345–1359.
- 584 Park, J. & Newman, M.E.J. (2004). Statistical mechanics of networks. *Physical Review E*, 70, 066117.
- 585 Pedersen, E.J., Thompson, P.L., Ball, R.A., Fortin, M.-J., Gouhier, T.C., Link, H., et al. (2017). Signatures of
586 the collapse and incipient recovery of an overexploited marine ecosystem. *Royal Society Open Science*,
587 4, 170215.

- 588 Perretti, C.T., Munch, S.B. & Sugihara, G. (2013). Model-free forecasting outperforms the correct
589 mechanistic model for simulated and experimental data. *Proceedings of the National Academy of*
590 *Sciences*, 110, 5253–5257.
- 591 Pires, M.M., Koch, P.L., Fariña, R.A., de Aguiar, M.A.M., dos Reis, S.F. & Guimarães, P.R. (2015).
592 Pleistocene megafaunal interaction networks became more vulnerable after human arrival.
593 *Proceedings of the Royal Society B: Biological Sciences*, 282, 20151367.
- 594 Poelen, J.H., Simons, J.D. & Mungall, C.J. (2014). Global biotic interactions: An open infrastructure to
595 share and analyze species-interaction datasets. *Ecological Informatics*, 24, 148–159.
- 596 Poisot, T., Belisle, Z., Hoebelke, L., Stock, M. & Szefer, P. (2019). EcologicalNetworks.jl - analysing
597 ecological networks. *Ecography*.
- 598 Poisot, T., Bergeron, G., Cazelles, K., Dallas, T., Gravel, D., MacDonald, A., *et al.* (2021a). Global
599 knowledge gaps in species interaction networks data. *Journal of Biogeography*, n/a.
- 600 Poisot, T., Cirtwill, A.R., Cazelles, K., Gravel, D., Fortin, M.-J. & Stouffer, D.B. (2016). The structure of
601 probabilistic networks. *Methods in Ecology and Evolution*, 7, 303–312.
- 602 Poisot, T., Ouellet, M.-A., Mollentze, N., Farrell, M.J., Becker, D.J., Albery, G.F., *et al.* (2021b). *Imputing the*
603 *mammalian virome with linear filtering and singular value decomposition*. Available at:
604 <http://arxiv.org/abs/2105.14973>. Last accessed.
- 605 Poisot, T. & Stouffer, D.B. (2018). Interactions retain the co-phylogenetic matching that communities lost.
606 *Oikos*, 127, 230–238.
- 607 Poisot, T., Stouffer, D.B. & Gravel, D. (2015). Beyond species: Why ecological interaction networks vary
608 through space and time. *Oikos*, 124, 243–251.
- 609 Price, P.W. (2003). *Macroevolutionary theory on macroecological patterns*. Cambridge University Press.
- 610 Ray, J.C., Grimm, J. & Olive, A. (2021). The biodiversity crisis in Canada: Failures and challenges of
611 federal and sub-national strategic and legal frameworks. *FACETS*, 6, 1044–1068.
- 612 Reeve, R., Leinster, T., Cobbolt, C.A., Thompson, J., Brummitt, N., Mitchell, S.N., *et al.* (2016). *How to*
613 *partition diversity*. Available at: <http://arxiv.org/abs/1404.6520>. Last accessed.
- 614 Rosado, B.H.P., Dias, A. & de Mattos, E. (2013). Going Back to Basics: Importance of Ecophysiology when
615 Choosing Functional Traits for Studying Communities and Ecosystems. *Natureza & conservação*

- 616 *revista brasileira de conservaç~ao da natureza*, 11, 15–22.
- 617 Runghen, R., Stouffer, D.B. & Dalla Riva, G.V. (2021). Exploiting node metadata to predict interactions in
618 large networks using graph embedding and neural networks.
- 619 Schoener, T.W. (1989). Food webs from the small to the large. *Ecology*, 70, 1559–1589.
- 620 Shlens, J. (2014). *A Tutorial on Principal Component Analysis*. Available at:
621 <http://arxiv.org/abs/1404.1100>. Last accessed.
- 622 Solís-Lemus, C., Bastide, P. & Ané, C. (2017). PhyloNetworks: A Package for Phylogenetic Networks.
623 *Molecular Biology and Evolution*, 34, 3292–3298.
- 624 Stock, M. (2021). Pairwise learning for predicting pollination interactions based on traits and phylogeny.
625 *Ecological Modelling*, 14.
- 626 Stouffer, D.B., Sales-Pardo, M., Sirer, M.I. & Bascompte, J. (2012). Evolutionary Conservation of Species'
627 Roles in Food Webs. *Science*, 335, 1489–1492.
- 628 Strong, J.S. & Leroux, S.J. (2014). Impact of Non-Native Terrestrial Mammals on the Structure of the
629 Terrestrial Mammal Food Web of Newfoundland, Canada. *PLOS ONE*, 9, e106264.
- 630 Strydom, T., Catchen, M.D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., et al. (2021a). A
631 roadmap towards predicting species interaction networks (across space and time). *Philosophical
632 Transactions of the Royal Society B: Biological Sciences*, 376, 20210063.
- 633 Strydom, T., Dalla Riva, G.V. & Poisot, T. (2021b). SVD Entropy Reveals the High Complexity of Ecological
634 Networks. *Frontiers in Ecology and Evolution*, 9.
- 635 Thessen, A.E. & Parr, C.S. (2014). Knowledge extraction and semantic annotation of text from the
636 encyclopedia of life. *PloS one*, 9, e89550.
- 637 Torrey, L. & Shavlik, J. (2010). Transfer learning. In: *Handbook of research on machine learning
638 applications and trends: Algorithms, methods, and techniques*. IGI global, pp. 242–264.
- 639 Trøjelsgaard, K. & Olesen, J.M. (2016). Ecological networks in motion: Micro- and macroscopic variability
640 across scales. *Functional Ecology*, 30, 1926–1935.
- 641 Turak, E., Brazill-Boast, J., Cooney, T., Drielsma, M., DelaCruz, J., Dunkerley, G., et al. (2017). Using the
642 essential biodiversity variables framework to measure biodiversity change at national scale. *Biological*

- 643 *Conservation*, SI:measures of biodiversity, 213, 264–271.
- 644 Upham, N.S., Esselstyn, J.A. & Jetz, W. (2019). Inferring the mammal tree: Species-level sets of
645 phylogenies for questions in ecology, evolution, and conservation. *PLOS Biology*, 17, e3000494.
- 646 van der Hoorn, P., Lippner, G. & Krioukov, D. (2018). Sparse Maximum-Entropy Random Graphs with a
647 Given Power-Law Degree Distribution. *Journal of Statistical Physics*, 173, 806–844.
- 648 Vermaat, J.E., Dunne, J.A. & Gilbert, A.J. (2009). Major dimensions in food-web structure properties.
649 *Ecology*, 90, 278–282.
- 650 Wood, S.A., Russell, R., Hanson, D., Williams, R.J. & Dunne, J.A. (2015). Effects of spatial scale of
651 sampling on food web structure. *Ecology and Evolution*, 5, 3769–3782.
- 652 Xing, S. & Fayle, T.M. (2021). The rise of ecological network meta-analyses: Problems and prospects.
653 *Global Ecology and Conservation*, 30, e01805.
- 654 Yeakel, J.D., Pires, M.M., Rudolf, L., Dominy, N.J., Koch, P.L., Guimarães, P.R., *et al.* (2014). Collapse of an
655 ecological network in Ancient Egypt. *PNAS*, 111, 14472–14477.
- 656 Youden, W.J. (1950). Index for rating diagnostic tests. *Cancer*, 3, 32–35.
- 657 Young, J.-G., Cantwell, G.T. & Newman, M.E.J. (2021). Bayesian inference of network structure from
658 unreliable data. *Journal of Complex Networks*, 8.
- 659 Young, S.J. & Scheinerman, E.R. (2007). Random Dot Product Graph Models for Social Networks. In:
660 *Algorithms and Models for the Web-Graph*, Lecture Notes in Computer Science (eds. Bonato, A. &
661 Chung, F.R.K.). Springer, Berlin, Heidelberg, pp. 138–149.
- 662 Zhu, M. & Ghodsi, A. (2006). Automatic dimensionality selection from the scree plot via the use of profile
663 likelihood. *Computational Statistics & Data Analysis*, 51, 918–930.

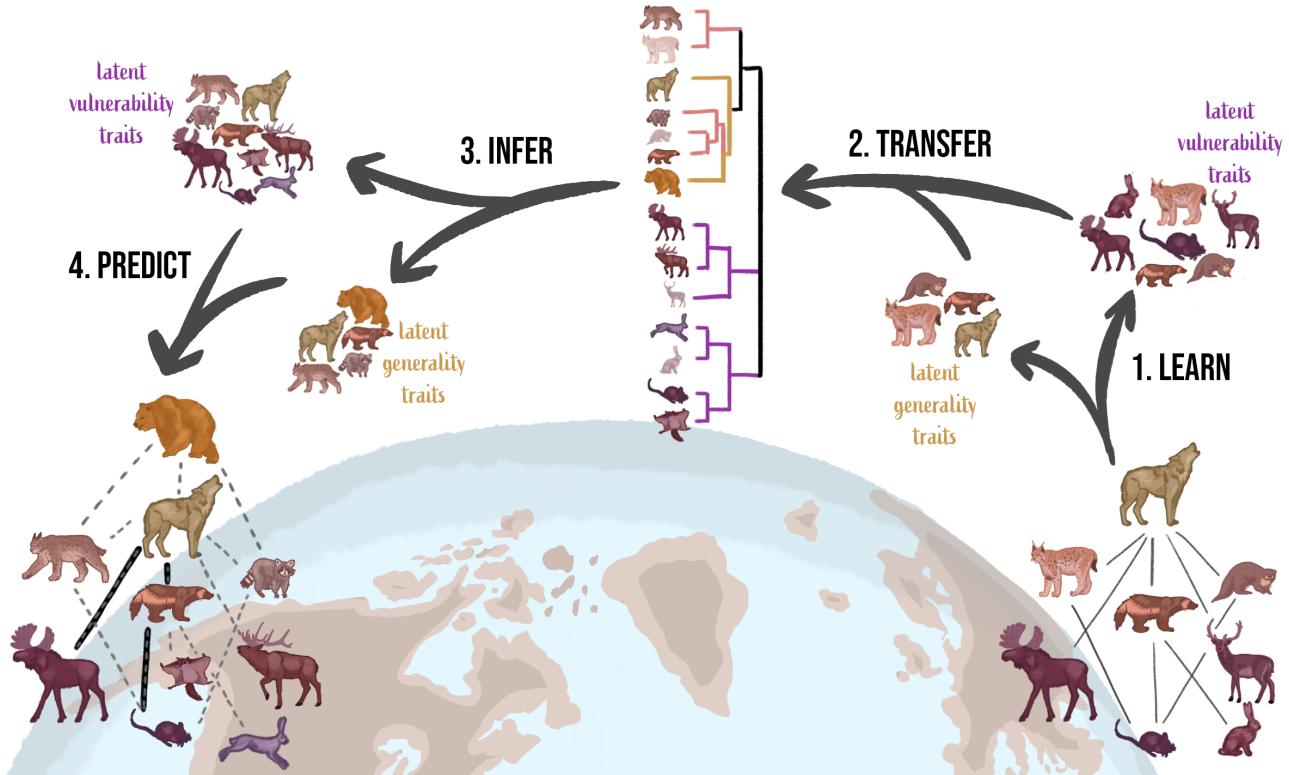


Figure 1: Overview of the phylogenetic transfer learning (and prediction) of species interaction networks. Starting from an initial, known, network, we learn its representation through a graph embedding step (here, a truncated Singular Value Decomposition; Step 1), yielding a series of latent traits (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).

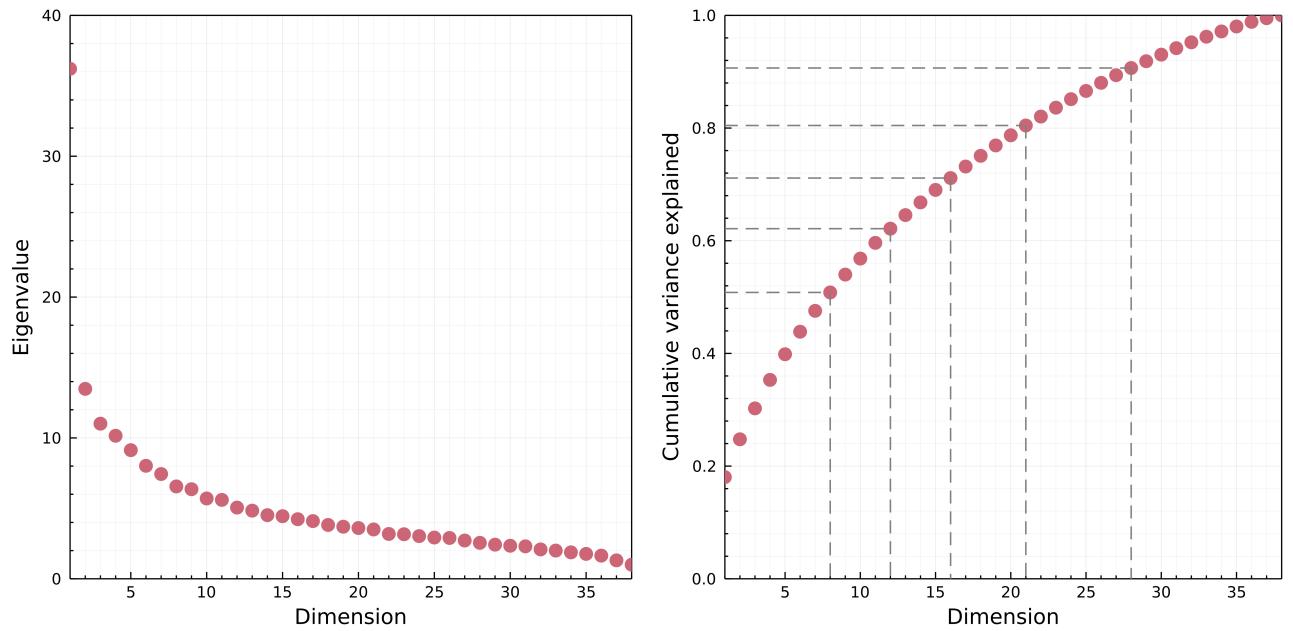


Figure 2: Left: representation of the scree plot of the singular values from the t-SVD on the European metaweb. The scree plot shows no obvious drop in the singular values that may be leveraged to automatically detect a minimal dimension for embedding, after e.g. Zhu & Ghodsi (2006). Right: cumulative 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.

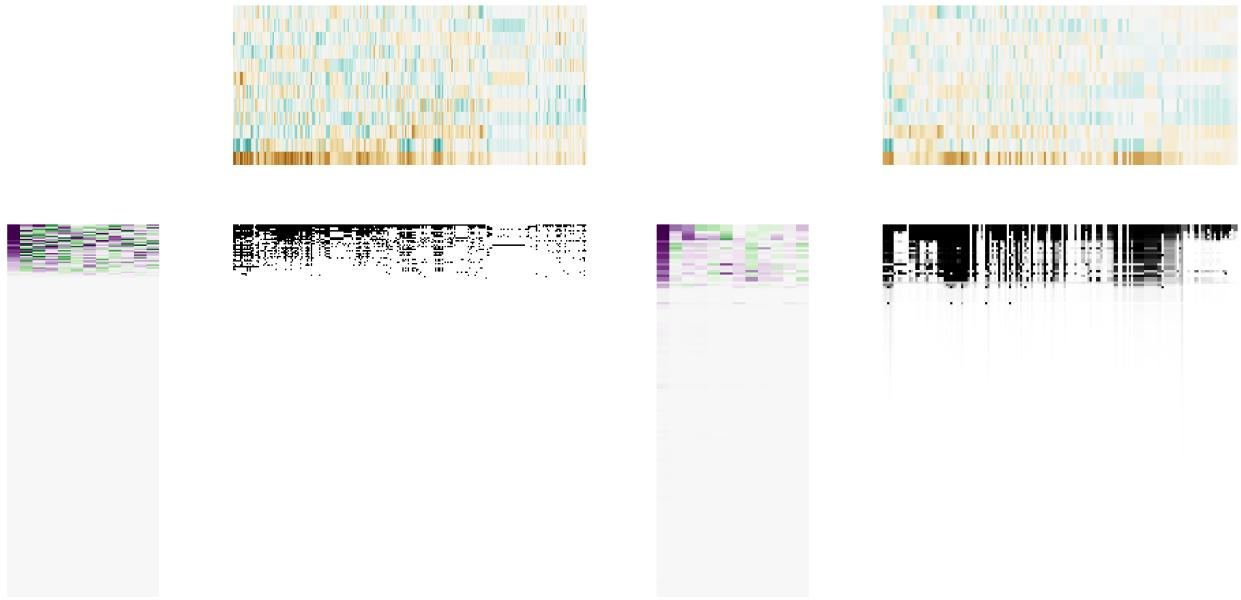


Figure 3: Visual representation of the left (green/purple) and right (green/brown) subspaces, alongside the adjacency matrix of the food web they encode (greyscale). The European metaweb is on the left, and the imputed Canadian metaweb (before data inflation) on the right. This figure illustrates how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

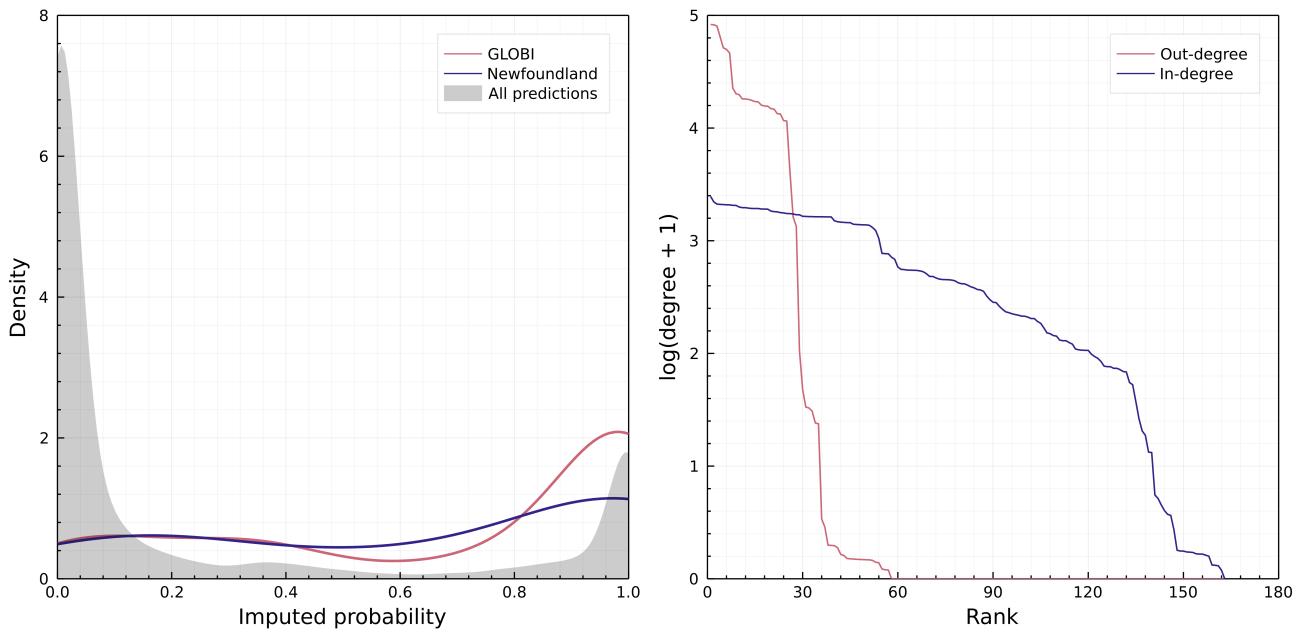


Figure 4: Left, comparison of the probabilities of interactions assigned by the model to all interactions (grey curve), the subset of interactions found in GLOBI (red), and in the Strong & Leroux (2014) Newfoundland dataset (blue). The model recovers more interactions with a low probability compared 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.

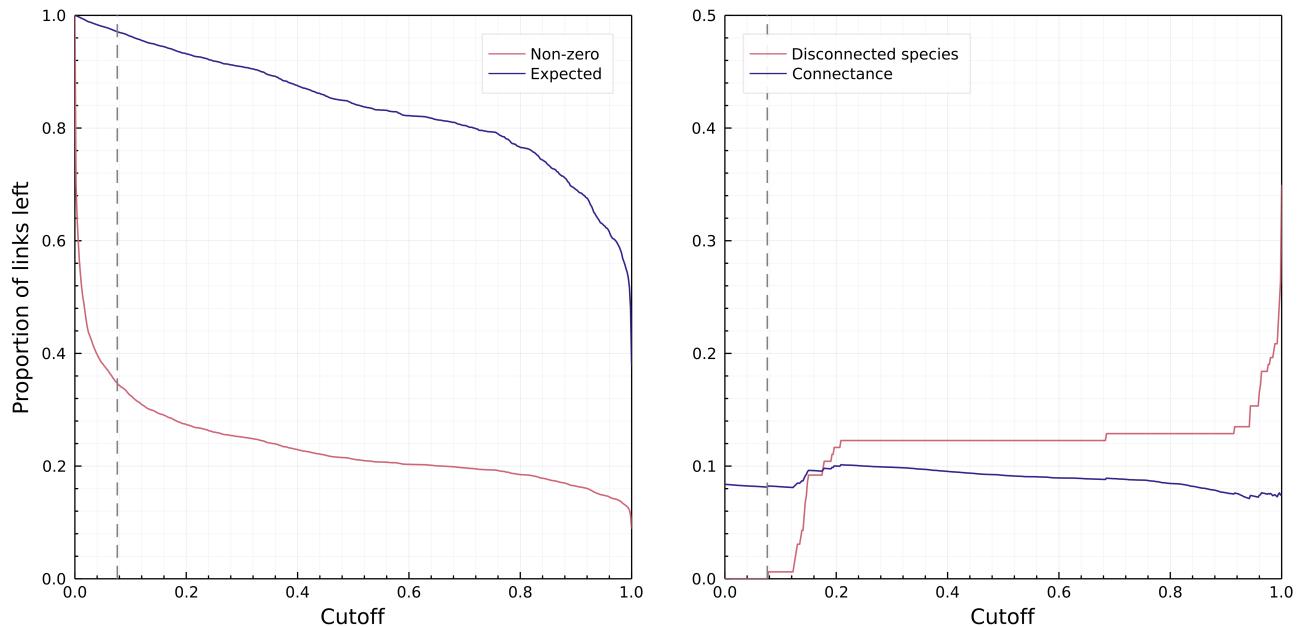


Figure 5: Left: effect of varying the cutoff for probabilities to be considered non-zero on the number of unique links and on \hat{L} , the probabilistic estimate of the number of links assuming that all interactions are independent. Right: effect of varying the cutoff on the number of disconnected species, 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.

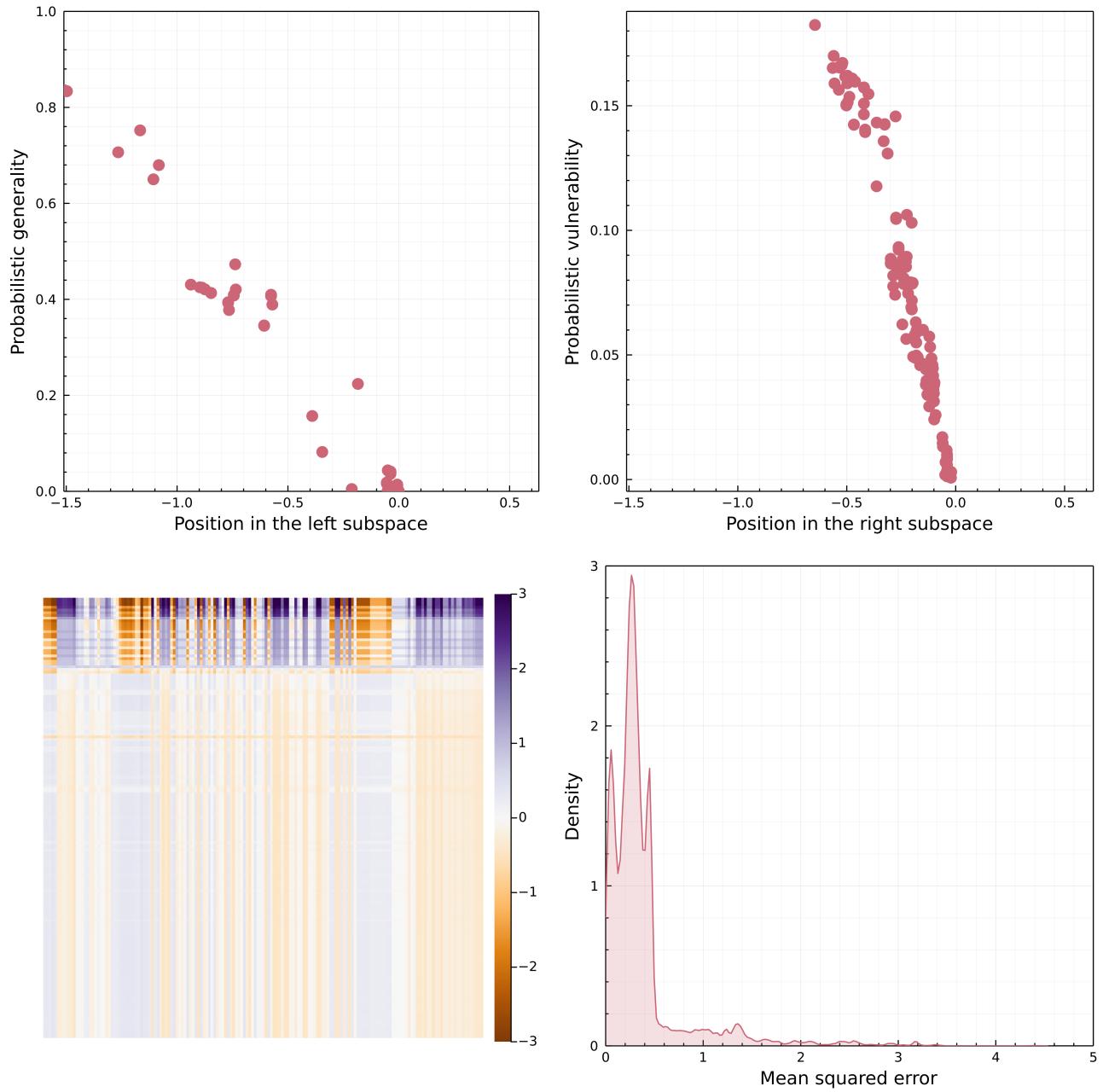


Figure 6: Top: biological significance of the first dimension. Left: there is a linear relationship between the values on the first dimension of the left subspace and the generality, *i.e.* the relative number of preys, *sensu* Schoener (1989). Species with a value of 0 in this subspace are at the bottom-most 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-score 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.