# S.M.1 - SVD does not overfit on the European network

In order to ensure that the creation of the RDPG on the European network does not lead to overfitting, we performed the following numerical experiment. First, we extracted $\mathcal{L}$ and $\mathcal{R}$, the left and right subspace of the entire network, at rank 12. Then, for every number n of interactions between 10 and $\mathrm{links}(M) - 1$ (where M is the European metaweb), we define m as a network in which n interactions have been randomly removed. We then define $\updownarrow$ and ∇ as the left and right subspaces coming from the rank-12 RDPG applied to this network, and compare the original network M to the one that was reconstructed after thresholding $\updownarrow\nabla$ by picking the cutoff that maximizes Youden's J measure.

This experiment allows masuring the response of various measures of fit of the binary classifier to incomplete sampling. We are specifically interested in (i) the ability of RDPG to identify removed interactions, (ii) the ability of RDPG to function as a performant classifier in the presence of uncertainty in the original data, and (iii) the ability of RDPG to reconstruct biologically realistic data when interactions are withheld.

## RDPG recovers withheld interactions

When removing up to 40% of interactions in the European metaweb, the RDPG was able to correctly recover 80% of these interactions:
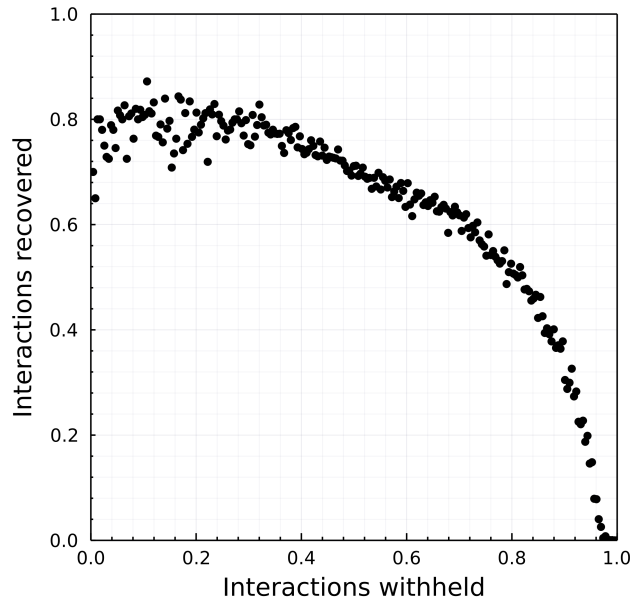
The stochasticity in the proportion of recovered interactions is larger when a

small number of interactions are withheld, which makes sense as the *number* of interactions is far smaller (compared to the overall network size).

It is interesting to note that the threshold "adapts" to the amount of missing information - the dashed line corresponds to the threshold we used in the manuscript.
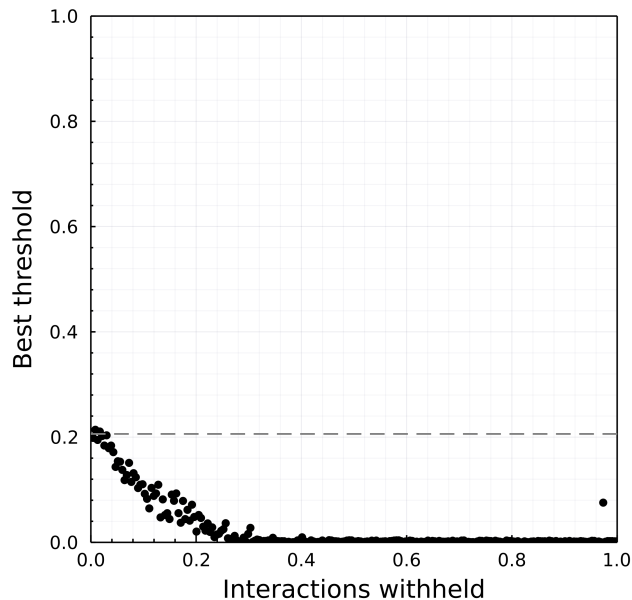

The important consequence of this result is that training the RDPG on a sub-sample of the network would result in a lower threshold, thereby creating more false positives when applied to the new data.

RDPG yields an accurate classifier

More important than the recovery of removed interaction is the fact that the classifier should not create too much false positives. One measure to assess this is the area under the receiving operator characteristic curve, or ROC-AUC. By this measure, the RDPG remains an excellent classifier even if 50% of interactions are withheld.
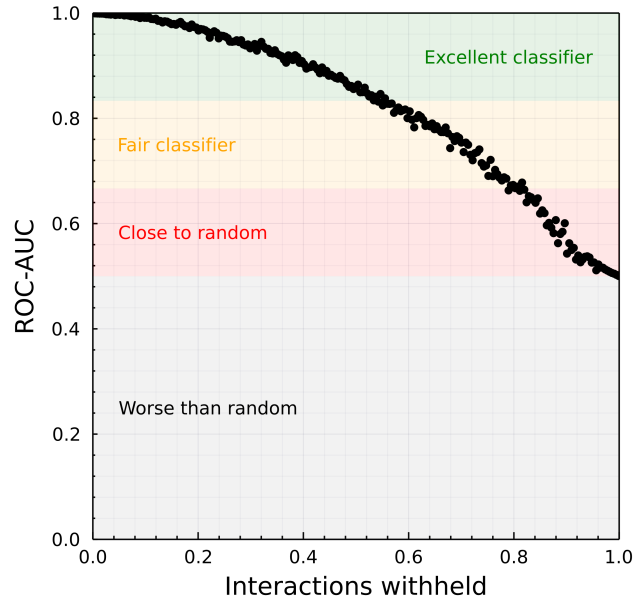
The overall agreement between a classifier and the actual data can be measured by Cohen's κ, which gives a similar result.
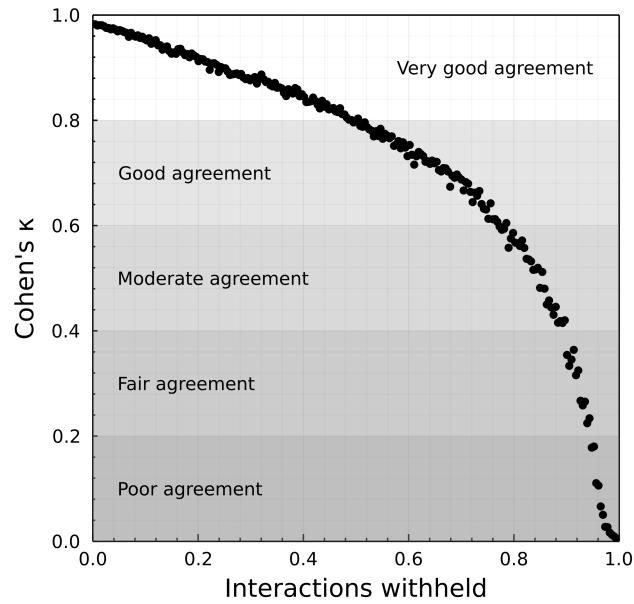
These two diagnostic figures reveal that, although we used a probably exhaustive list of interactions to do the initial RDPG, there are chances that the approach would work on less complete datasets.

RDPG creates ecologically realistic networks

In this section, we present the relationship between the empirical measure of the network structure (dashed line) and the reconstructed estimate based on RDPG after the optimal threshold has been applied. We focus on connectance (for its broad relevant to food web structure) first:
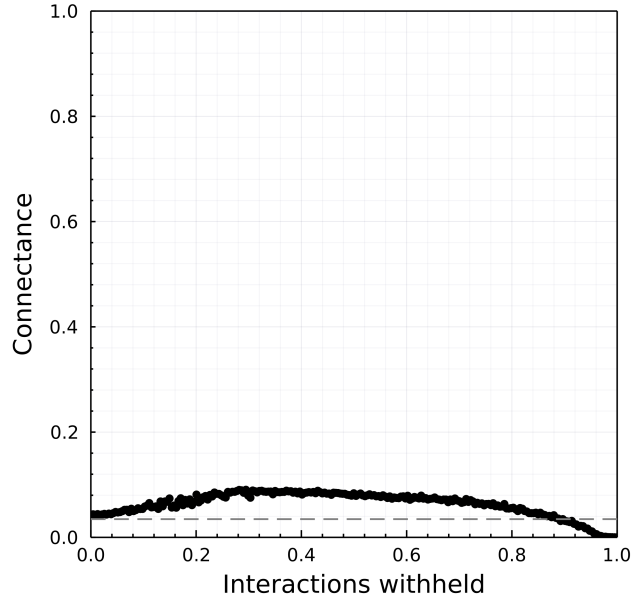
Connectance increases slightly when initial information is incomplete, but saturates at a value of around 0.12 – this is still within the bounds of connectances expected for food webs.

Next, we look at the ratio between direct competition $(a \to (b, c))$ and apparent competition $((a, b) \to c)$ motifs, as motifs are known to be conserved blocks in food webs:
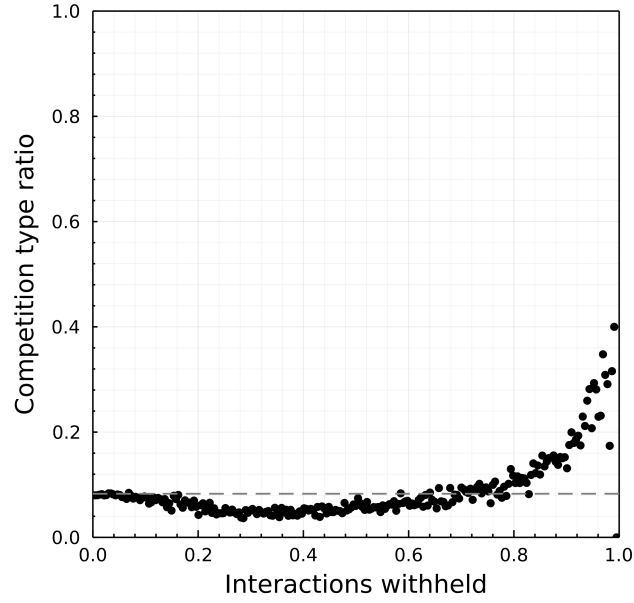
This ratio remains close to the real one up until 75% of initial interactions are lost.
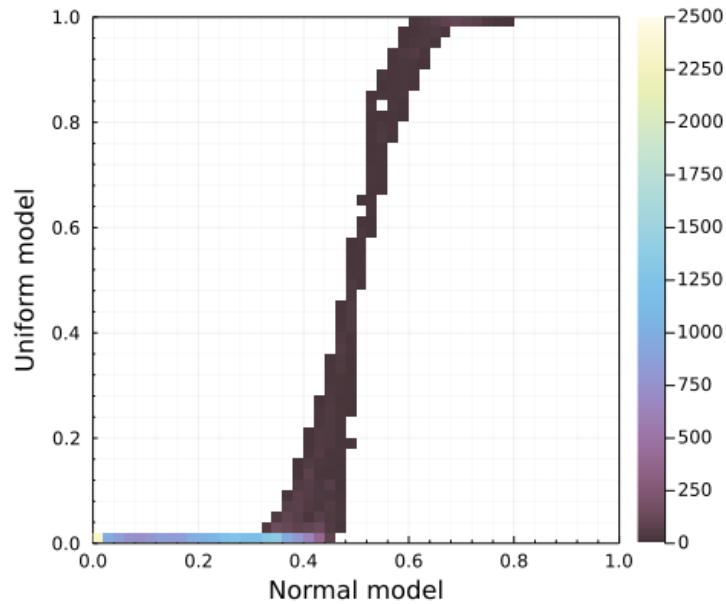
Consequences

Based on these results, applying RDPG on the entire European network is reasonable, especially since the threshold is adapting to the amount of withheld interactions. Interestingly, the RDPG remains an excellent binary classifier even in the face of strong data deficiencies, which suggests that our framework can be used even in the absence of a complete metaweb.

S.M.2 - the Normal model of latent variable evolution over-predicts

In this appendix, we compare the raw predictions made by the Normal and Uniform models of latent variable evolution. The Normal model was created by (i) getting the averag μ of the simulated values for each species/variable combination, and (ii) estimating the standard deviation as $(\mu + c - \mu - c)/3.92$, where c is one half of the 95% confidence interval around μ.

As can be seen on the following figure, the Normal model tends to assign high probabilities (up to $p \approx 0.4$) for interactions that the Uniform model essentially rules out:


This can lead to severe over-estimation of the number of interactions. In fact, the consequences of using a Normal model are obvious from looking at the adjacency matrices below: most of the interactions are predicted between species that occupy the lower trophic level, and are ecologically unrealistic.
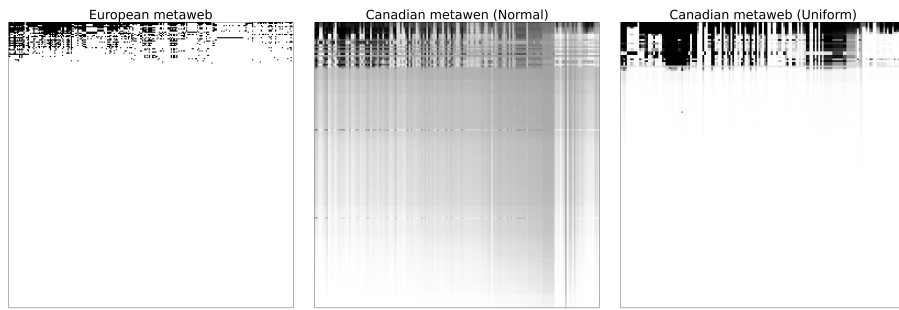
For this reason, we only use predictions from the Uniform model in the main text.

S.M.3 - RDPG reconstructed networks have diverse structures

In this appendix, we check that the networks reconstructed from the RDPG do keep a variety of structural components, especially when selecting small species pools from within them. In order to do so, we induced 400 random subgraphs containing within 30 and 70 species, both from the Canadian and European metaweb. For each of these subgraphs, we measured eight variables: the mean and standard deviation of trophic levels, the standard deviation of degree (total, in, and out), and the proportion of top, intermediate, and basal species. We selected a random subset of 300 rows from the network-property matrix to fit a Principal Component Analysis projection matrix (W), which we then used to project all networks into the space formed by the first two principal components.
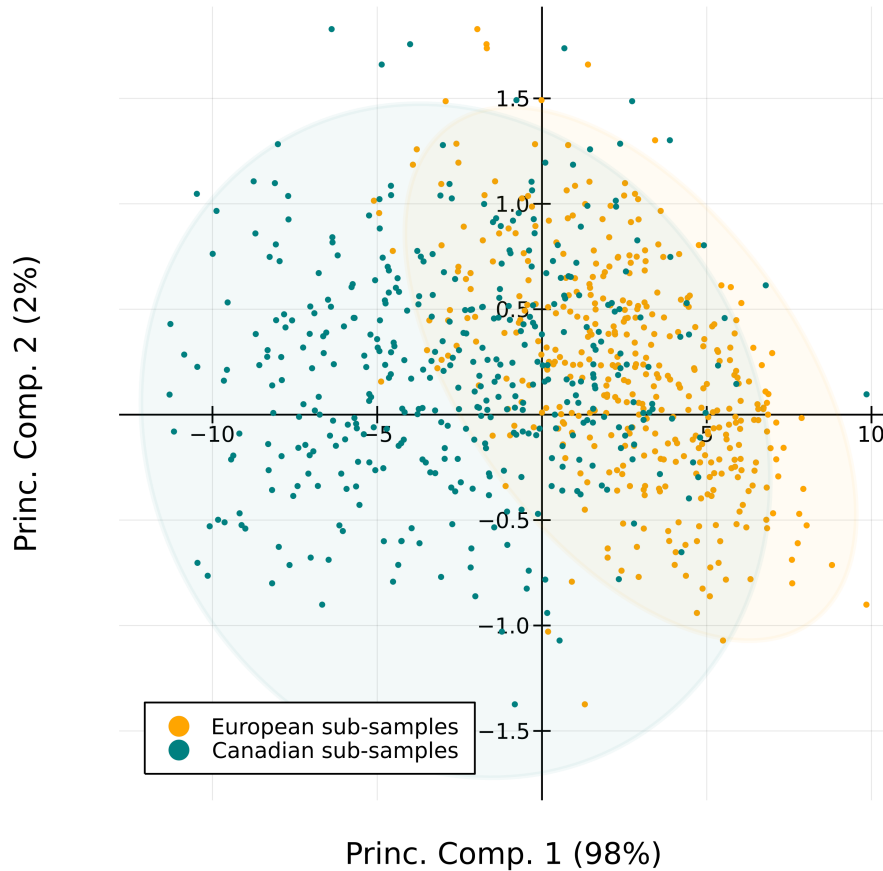
The first axis (explaining most variance) was strongly correlated to the standard deviation of the number of preys (-0.71), and the second axis to the standard

deviation in the number of predators (-0.95). These results match the conclusions in main text, namely that the first dimensions of network embedding capture the degree distribution.

Two things are important to note on this representation; each point is an induced sub-graph, and the ellipses are the 95% confidence interval around the points. First, there is some variations *within* a group (Europe *v.* Canada); second, the two groups do not fully overlap. This suggests that not only the sub-samples of the Canadian metaweb are not equivalent to the sub-samples of the European metaweb (*i.e.* the two networks have structural differences), realizations (here in the form of random local species pools) of the Canadian metaweb also show some variability; in short, reconstructing a metaweb using a RDPG will not result in homogeneous local networks, and may therefore be suitable for lower-scale predictions.