

# Maximum entropy modelling of food-web structure

Francis Banville<sup>1,2,3</sup> Dominique Gravel<sup>2,3</sup> Timothée Poisot<sup>1,3</sup>

<sup>1</sup> Université de Montréal <sup>2</sup> Université de Sherbrooke <sup>3</sup> Quebec Centre for Biodiversity Science

## Correspondance to:

Francis Banville — francis.banville@umontreal.ca

**Abstract:** The principle of maximum entropy is a rigorous mathematical method of finding constrained probability distributions that has been proven useful in many ecological modelling problems. However, despite its broad application in graph and network theory, it has seldom been used to model ecological networks. Here we show how the (joint) degree distribution of maximum entropy can be directly derived using the number of species and the number of links in terrestrial and aquatic food webs. We also present a heuristic and flexible approach of finding the network of maximum entropy based on simulating annealing and SVD entropy. We built two of these network-level models using constraints given by the connectance and the joint degree sequence, respectively. All maximum entropy models were compared against open access food-web data and null and neutral models commonly used in network ecology. We found that the maximum entropy network model constrained by the joint degree sequence was a good predictor of food-web structure, including nestedness and motifs distribution. Overall, our results suggest that many properties of ecological networks are mainly driven by the joint degree distribution and statistical phenomena.

## **1 Introduction**

2 Statistical and mathematical models can help fill many gaps in our knowledge about species in-  
3 teractions. Predictive and null models are two complementary types of models that have been  
4 developed in network ecology for this purpose. On one hand, predictive models can partially  
5 alleviate the Eltonian shortfall, which describes our current lack of knowledge on food webs  
6 and other ecological networks (Hortal et al. 2015). A variety of such models have recently  
7 been developed using machine learning and other statistical tools, most of which are presented  
8 in Strydom et al. (2021). On the other hand, null models help us identify potential ecological  
9 mechanisms that drive species interactions. They do so by comparing empirical data with an  
10 unbiased distribution of measures generated using a set of rules that exclude the mechanism of  
11 interest (Fortuna and Bascompte 2006; Delmas et al. 2019). Both types of models are frequently  
12 topological, i.e. they often predict the adjacency matrix or specific measures of network struc-  
13 ture without taking into account species' identity. According to Strydom et al. (2021), these  
14 topological models could be used to make better predictions of pairwise species interactions by  
15 constraining the space of feasible networks.

16 The principle of maximum entropy (MaxEnt) is a statistical and topological model that can be  
17 used for both of these purposes, i.e. to make predictions of network structure and to better un-  
18 derstand processes shaping ecological networks. This mathematical method, briefly presented  
19 in Box 1, has been used in a wide range of disciplines, from thermodynamics to chemistry and  
20 biology (Martyushev and Seleznev 2006). It has also been proven useful in ecology, e.g. in  
21 species distribution models (Phillips, Anderson, and Schapire 2006) and macroecological mod-  
22 els (Harte et al. 2008; Harte and Newman 2014). As discussed in Box 1, maximizing a measure  
23 of entropy ensures that the derived probability distributions are unique and least biased under  
24 the set of constraints used. These constraints are built using state variables, i.e. variables that  
25 represent the macrostate of the system. The challenge is to find the set of state variables that best  
26 represent natural systems and to translate them into appropriate statistical constraints. Having  
27 a validated maximum entropy model for the system at hand allows us to make rigorous predic-  
28 tions using a minimal amount of data, as well as helping us describe the most important factors  
29 driving that system.

30 Despite its extensive use in graph and network theories (e.g., Park and Newman 2004; van der  
31 Hoorn, Lippner, and Krioukov 2018), MaxEnt has in comparison been little used in network  
32 ecology. Like other real system networks, ecological networks are represented mathematically  
33 as graphs. However, the very nature of ecological networks (directed simple graphs frequently  
34 having self-loops) makes the mathematical optimization of maximum entropy graph models  
35 more complicated than with many other types of (non-ecological) networks. MaxEnt has nev-  
36 ertheless been used to predict the degree distribution of bipartite ecological networks from the  
37 number of species and the number of interactions (Williams 2011) and to predict interaction  
38 strengths between species pairs using their relative abundances within an optimal transportation  
39 theory regularized with information entropy (Stock, Poisot, and De Baets 2021). However, to  
40 the best of our knowledge, MaxEnt has never been used to predict food-web structure directly,  
41 even though food webs are among the most documented and widespread ecological networks.

42 In this contribution, we used two complementary approaches to predict the structure of food  
43 webs using the principle of maximum entropy. We then compared our predictions against em-  
44 pirical data and null and neutral models commonly used in network ecology. The first approach  
45 consists in deriving constrained probability distributions of given network properties directly.  
46 We derived the joint degree distribution (a probability distribution) of maximum entropy using  
47 only the number of species  $S$  and the number of interactions  $L$  as state variables. Then, we  
48 predicted the degree distribution of maximum entropy directly from the joint degree distribu-  
49 tion since the first is the sum of the marginal distributions of the second (a species' degree is  
50 the sum of its in and out-degrees). Because of the scarcity of empirical data on the number of  
51 interactions in ecological networks, we present a method to predict  $L$  from  $S$  (Box 2), thus al-  
52 lowing the prediction of the joint degree distribution from  $S$  solely. In turn, the second approach  
53 consists in finding, using different constraints, the adjacency matrix of maximum entropy from  
54 which network properties can be measured. To do so, we used a flexible and heuristic approach  
55 based on simulated annealing to find networks *close* to maximum entropy. As discussed above,  
56 our choice of algorithm stands from the very nature of food webs (i.e. simple directed networks  
57 allowing self-loops) that makes the analytical derivation of a maximum entropy graph model  
58 difficult. We first built our type I MaxEnt network model constrained by the connectance of

59 the network (i.e. the ratio  $L/S^2$ ). A comparison of this model against empirical data indicated  
60 that connectance alone was not sufficient to predict many aspects of network structure. For this  
61 reason, we built our type II MaxEnt network model, which instead uses the whole joint degree  
62 sequence as a constraint. Overall, we found that this second model was much better at predicting  
63 food-web structure than the one based on connectance.

#### 64 **Box 1 - The principle of maximum entropy: A primer for ecologists**

65 The principle of maximum entropy is a mathematical method of finding probability distribu-  
66 tions, strongly rooted in statistical mechanics and information theory (Jaynes 1957a, 1957b;  
67 Harremoës and Topsøe 2001). Starting from a set of constraints given by prior knowledge of a  
68 system (i.e. what we call state variables), this method helps us find least-biased probability dis-  
69 tributions subject to the constraints. These probability distributions are guaranteed to be unique  
70 given our prior knowledge and represent the most we can say about a system without making  
71 more assumptions. For example, if the only thing we know about a biological community is  
72 its average number of individuals per species, the least-biased inference we could make on its  
73 species abundance distribution is the exponential distribution (Frank and Smith 2011; Harte  
74 and Newman 2014). However, this does not imply that this distribution will be the best fit to  
75 empirical data. The challenge is to find the right set of constraints that would best reproduce  
76 distributions found in nature.

77 Entropy measures the amount of information given by the outcome of a random variable. Many  
78 measures of entropy have been developed in physics (Beck 2009), but only a fraction of them  
79 could be used as an optimization measure with the principle of maximum entropy. According  
80 to Beck (2009) and Khinchin (2013), a measure of entropy  $H$  should satisfy four properties  
81 in the discrete case: (1) it should be a function of a probability distribution  $p(n)$  only; (2) it  
82 should be maximized when  $p(n)$  is uniform; (3) it should not be influenced by outcomes with a  
83 null probability; and (4) it should be independent of the order of information acquisition. The  
84 Shannon's entropy (Shannon 1948)

$$H = - \sum_n p(n) \log p(n) \quad (1)$$

85 satisfies all of these properties. Finding the probability distribution  $p(n)$  that maximizes  $H$  under  
 86 a set of  $m$  constraints  $g$  can be done using the method of Lagrange multipliers. These constraints  
 87 could include one or many properties of the probability distribution (e.g., its mean, variance, and  
 88 range). However, the normalization constraint always need to be included in  $g$  in order to make  
 89 sure that  $p(n)$  sums to 1. The objective is then to find the values of the Lagrange multipliers  $\lambda_i$   
 90 that optimize a function  $F$ :

$$F = H - \sum_{i=1}^m \lambda_i(g_i - c_i), \quad (2)$$

91 where  $g_i$  is the mathematical formulation of the constraint  $i$  and  $c_i$ , its value. Note that  $F$  is just  
 92 Shannon's entropy to which we added terms that each sums to zero ( $g_i = c_i$ ).  $F$  is maximized by  
 93 setting to 0 its partial derivative with respect to  $p(n)$ . We will show how this can be done when  
 94 we derive the joint degree distribution analytically from the number of species and the number  
 95 of interactions in food webs.

96 In this contribution, we also use the SVD entropy as a measure of entropy, which is an applica-  
 97 tion of Shannon's entropy to the relative non-zero singular values of a truncated singular value  
 98 decomposition (t-SVD; Strydom, Dalla Riva, and Poisot 2021) of a food web's Boolean ad-  
 99 jacency matrix. This measure also satisfies all four properties above-mentioned, while being a  
 100 proper measure of the internal complexity of food webs (Strydom, Dalla Riva, and Poisot 2021).

101 We measured SVD entropy as follows:

$$J = - \sum_{i=1}^R s_i \log s_i, \quad (3)$$

102 where  $s_i$  are the relative singular values ( $s_i = \sigma_i / \sum_{i=1}^R \sigma_i$ , where  $\sigma_i$  are the singular values).  
 103 Following Strydom, Dalla Riva, and Poisot (2021), we standardized this measure with the rank  $R$   
 104 of the matrix (i.e.  $J / \ln(R)$ ) to account for the difference in dimensions between networks (*sensu*  
 105 Pielou's evenness; Pielou 1975). In a following section, we will show how SVD entropy can be

106 used to predict a network of maximum entropy (i.e. of maximum complexity) heuristically.

## 107 Testing MaxEnt models

### 108 Data

109 We tested our MaxEnt models (both approaches) against open food-web data queried from three  
110 different sources and integrated into what we call our *complete dataset*. First, all food webs  
111 archived on `mangal.io` (Poisot et al. 2016; Banville, Vissault, and Poisot 2021) were directly  
112 queried from the database ( $n = 235$ ). Most ecological networks archived on Mangal are multi-  
113 layer networks, i.e. networks that describe different types of interactions. We kept all networks  
114 whose interactions were mainly of predation and herbivory types, and removed the largest net-  
115 work ( $S = 714$ ) for computational efficiency reasons. Then, to this set we added food webs  
116 from two different sources: the New-Zealand dataset ( $n = 21$ ; Pomeranz et al. 2018) and the  
117 Tuesday lake dataset ( $n = 2$ ; Cohen, Jonsson, and Carpenter 2003). Of these two datasets,  
118 networks had data on species' relative abundances that were used in the neutral model presented  
119 in a following subsection. These networks are part of what we call our *abundance dataset*, which  
120 is a subset of our complete dataset.

121 All code and data to reproduce this article are available at the Open Science Framework. Data  
122 cleaning, simulations and analyses were conducted in Julia v1.5.4.

### 123 Null models (types I and II)

124 Our maximum entropy network models (second approach only) were compared with two topo-  
125 logical null models. The first is the type I null model of Fortuna and Bascompte (2006), in which  
126 the probability that a species  $i$  predares on another species  $j$  is given by

$$p(i \rightarrow j) = \frac{L}{S^2}. \quad (4)$$

127 The second is the type II null model of Bascompte et al. (2003), in which the probability of

128 interaction is instead given by

$$p(i \rightarrow j) = \frac{1}{2} \left( \frac{k_{in}(j)}{S} + \frac{k_{out}(i)}{S} \right), \quad (5)$$

129 where  $k_{in}$  and  $k_{out}$  are the in and out-degrees, respectively. The type I null model is based on  
130 connectance, whereas the type II null model is based on the joint degree sequence. Therefore,  
131 the type I and II topological null models correspond with our type I and II MaxEnt network  
132 models, respectively, since they use similar constraints.

133 We predicted both types of null networks for all empirical networks in our complete dataset  
134 ( $n = 257$ ). We converted all probabilistic networks to Boolean networks by generating 100  
135 random Boolean networks for each of these probabilistic webs. Then, we counted the number  
136 of times each interaction was sampled, and kept the  $L$  entries that were drawn the most amount  
137 of time, with  $L$  given by the number of interactions in each food web. This ensured that the  
138 resulting null networks had the same number of interactions as their empirical counterparts.

## 139 Neutral model

140 We also compared our MaxEnt network models with a neutral model of relative abundances, in  
141 which the probabilities of interaction are given by

$$p(i \rightarrow j) \propto \frac{n_i}{N} \times \frac{n_j}{N}, \quad (6)$$

142 where  $n_i$  and  $n_j$  are the abundances (or biomass) of both species, and  $N$  is the total abundance  
143 (or biomass) of all species in the network. We predicted neutral abundance matrices for all  
144 empirical networks in our abundance dataset ( $n = 19$ ), and converted these weighted matrices  
145 to Boolean networks using an approach analogue to the one we used for our null models.

<sup>146</sup> **Analytical models: Measures of maximum entropy**

<sup>147</sup> **Joint degree distribution**

<sup>148</sup> The joint degree distribution  $p(k_{in}, k_{out})$  is a joint discrete probability distribution describing  
<sup>149</sup> the probability that a species has  $k_{in}$  predators and  $k_{out}$  preys, with  $k_{in}$  and  $k_{out} \in [0, S]$ . Basal  
<sup>150</sup> species (e.g., plants) have a  $k_{out}$  of 0, whereas top predators have a  $k_{in}$  of 0. In contrast, the  
<sup>151</sup> maximum number of preys and predators a species can have is set by the number of species  $S$   
<sup>152</sup> in the food web. Here we show how the joint degree distribution of maximum entropy can be  
<sup>153</sup> obtained given knowledge of  $S$  and  $L$ .

<sup>154</sup> We want to maximize Shannon's entropy

$$H = - \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S p(k_{in}, k_{out}) \log p(k_{in}, k_{out}) \quad (7)$$

<sup>155</sup> subject to the following constraints:

$$g_1 = \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S p(k_{in}, k_{out}) = 1; \quad (8)$$

$$g_2 = \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S k_{in} p(k_{in}, k_{out}) = \langle k_{in} \rangle = \frac{L}{S}; \quad (9)$$

$$g_3 = \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S k_{out} p(k_{in}, k_{out}) = \langle k_{out} \rangle = \frac{L}{S}. \quad (10)$$

<sup>156</sup> The first constraint  $g_1$  is our normalizing constraint, whereas the other two ( $g_2$  and  $g_3$ ) fix the  
<sup>157</sup> average of the marginal distributions of  $k_{in}$  and  $k_{out}$  to the linkage density  $L/S$ . It is important  
<sup>158</sup> to notice that  $\langle k_{in} \rangle = \langle k_{out} \rangle$  because every edge is associated to a predator and a prey. Therefore,  
<sup>159</sup> without any further constraints, we expect the joint degree distribution of maximum entropy to  
<sup>160</sup> be a symmetric probability distribution with regards to  $k_{in}$  and  $k_{out}$ . However, this does not mean  
<sup>161</sup> that the joint degree *sequence* will be symmetric, since the joint degree sequence is essentially  
<sup>162</sup> a random realization of its probabilistic counterpart.

<sup>163</sup> The joint probability distribution of maximum entropy given these constraints is found using the  
<sup>164</sup> method of Lagrange multipliers. To do so, we seek to maximize the following expression:

$$F = H - \lambda_1(g_1 - 1) - \lambda_2\left(g_2 - \frac{L}{S}\right) - \lambda_3\left(g_3 - \frac{L}{S}\right), \quad (11)$$

<sup>165</sup> where  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  are the Lagrange multipliers. The probability distribution that maximizes  
<sup>166</sup> entropy is obtained by finding these values. As pointed out in Box 1,  $F$  is just Shannon's entropy  
<sup>167</sup> to which we added terms that each sums to zero (our constraints).  $F$  is maximized by setting to  
<sup>168</sup> 0 its partial derivative with respect to  $p(k_{in}, k_{out})$ . Because the derivative of a constant is zero,  
<sup>169</sup> this gives us:

$$\frac{\partial H}{\partial p(k_{in}, k_{out})} = \lambda_1 \frac{\partial g_1}{\partial p(k_{in}, k_{out})} + \lambda_2 \frac{\partial g_2}{\partial p(k_{in}, k_{out})} + \lambda_3 \frac{\partial g_3}{\partial p(k_{in}, k_{out})}. \quad (12)$$

<sup>170</sup> Evaluating the partial derivatives with respect to  $p(k_{in}, k_{out})$ , we obtain:

$$-\log p(k_{in}, k_{out}) - 1 = \lambda_1 + \lambda_2 k_{in} + \lambda_3 k_{out}. \quad (13)$$

<sup>171</sup> Then, solving eq. 13 for  $p(k_{in}, k_{out})$ , we obtain:

$$p(k_{in}, k_{out}) = \frac{e^{-\lambda_2 k_{in} - \lambda_3 k_{out}}}{Z}, \quad (14)$$

<sup>172</sup> where  $Z = e^{1+\lambda_1}$  is called the partition function. The partition function ensures that probabilities  
<sup>173</sup> sum to 1 (our normalization constraint). It can be expressed in terms of  $\lambda_2$  and  $\lambda_3$  as follows:

$$Z = \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S e^{-\lambda_2 k_{in} - \lambda_3 k_{out}}. \quad (15)$$

<sup>174</sup> After substituting  $p(k_{in}, k_{out})$  in eq. 9 and eq. 10, we get a nonlinear system of two equations and  
<sup>175</sup> two unknowns:

$$\frac{1}{Z} \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S k_{in} e^{-\lambda_2 k_{in} - \lambda_3 k_{out}} = \frac{L}{S}; \quad (16)$$

$$\frac{1}{Z} \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S k_{out} e^{-\lambda_2 k_{in} - \lambda_3 k_{out}} = \frac{L}{S}. \quad (17)$$

176 We solved eq. 16 and eq. 17 numerically using the Julia library JuMP.jl v0.21.8 (Dunning,  
 177 Huchette, and Lubin 2017). JuMP.jl supports nonlinear optimization problems by providing  
 178 exact second derivatives that increase the accuracy and performance of its solvers. The estimated  
 179 values of  $\lambda_2$  and  $\lambda_3$  can be substituted in eq. 14 to have a more workable expression for the joint  
 180 degree distribution.

181 We predicted the joint degree distribution of maximum entropy for each food web in our com-  
 182 plete dataset, i.e. using their numbers of species and numbers of interactions as state variables.  
 183 We then sampled one realization of the degree sequence for each network using the probabilities  
 184 given by the joint degree distribution. In fig. 1 (left panels), we show the relationship between  
 185  $k_{out}$  and  $k_{in}$  standardized by the number of species in their networks, for empirical and maximum  
 186 entropy joint degree distributions. We see that our model predicts higher values of generality  
 187 and vulnerability compared to empirical food webs (i.e. relative values of  $k_{out}$  and  $k_{in}$  closer to  
 188 1). However, plotting the difference between predicted and empirical values for each species  
 189 gives a different perspective. The right panel of fig. 1 presents these differences when species  
 190 are ordered by their total degree in their network (i.e. by the sum of their in and out-degrees).  
 191 Indeed, our predicted joint degree sequences have the same number of species as their empiri-  
 192 cal counterparts, but they are species agnostic; in other words, instead of predicting a value for  
 193 each species directly, we predicted the entire joint degree sequence without taking into account  
 194 species' identity. When we associate predictions and empirical data according to their rank in  
 195 total degrees, we see that species predicted to have a higher generality (number of preys)  
 196 generally have a lower vulnerability (number of preys) than what is observed (and conversely).  
 197 In fig. S1, we show how these differences change when species are instead ordered by their  
 198 out-degrees (left panel) and in-degrees (right panel), respectively.

199

[Figure 1 about here.]

200 We plotted the Kullback–Leibler divergence between in and out-degree distributions to compare  
 201 the symmetry of empirical and maximum entropy joint degree sequences fig. 2. As we expected,  
 202 our model predicted more similar in-degree and out-degree distributions than empirical data.  
 203 However, this difference decreased with connectance right panel of 2. Overall, this suggests that  
 204 other ecological constraints might be needed to account for the asymmetry of the joint degree  
 205 distribution, especially for networks with a lower connectance.

206

[Figure 2 about here.]

## 207 Degree distribution

208 The degree distribution  $p(k)$  represents the probability that a species has  $k$  interactions in a food  
 209 web, with  $k = k_{in} + k_{out}$ . It can thus be directly obtained from the joint degree distribution:

$$p(k) = \sum_{i=0}^k p(k_{in} = k - i, k_{out} = i).$$

210 In fig. S2, we show that the degree distribution of maximum entropy, given  $S$  and  $L$ , predicts  
 211 very low probabilities that a species will be isolated in its food web (i.e. having  $k = 0$ ). As  
 212 MacDonald, Banville, and Poisot (2020) pointed out, the size of food webs should at least be  
 213 of  $S - 1$  interactions, since a lower number would yield isolated species, i.e. species without  
 214 any predators or preys. Our results show that, under our purely information-theoretic model, the  
 215 probability that a species is isolated is quite high below this threshold. The expected proportion  
 216 of isolated species rapidly declines by orders of magnitude with increasing numbers of species  
 217 and interactions.

218 The degree distribution could also have been obtained directly using the principle of maximum  
 219 entropy, as discussed in Williams (2011). This gives the following distribution:

$$p(k) = \frac{e^{-\lambda_2 k}}{Z}, \quad (18)$$

220 with  $Z = \sum_{k=0}^S e^{-\lambda_2 k}$ .

221 This can be solved numerically using the constraint of the average degree  $\langle k \rangle = \frac{2L}{S}$  of a species.

222 Note that the mean degree is twice the value of the linkage density, because every link must be

223 counted twice when we add in and out-degrees together.

$$\frac{1}{Z} \sum_{k=0}^S k e^{-\lambda_2 k} = \frac{2L}{S} \quad (19)$$

224 The numerical solution is identical to the one we obtained using the joint degree distribution  
225 as an intermediate. Ecologists wanting to model a system without considering isolated species  
226 could simply change the lower limit of  $k$  to 1 and solve the resulting equation numerically.

227 In this section, we showed how important measures of food-web structure, namely the degree  
228 distribution and the joint degree distribution, could be derived with the principle of maximum  
229 entropy using minimal knowledge on a biological community. This type of models, although  
230 useful to make least-biased predictions on many network properties, can be hard to apply for  
231 other measures. Indeed, there are dozens of measures of network structure (Delmas et al. 2019)  
232 and many are not calculated with mathematical equations, but with algorithms. Moreover, the  
233 applicability of this method to empirical systems is limited by the state variables we can actu-  
234 ally measure and use. In the next section, we propose a more flexible method to predict many  
235 measures of network structure simultaneously, i.e. by finding networks of maximum entropy  
236 heuristically.

## 237 **Box 2 - Working with predicted numbers of interactions**

238 Our models need information on the number of species and the number of interactions. However,  
239 since the later is rarely estimated empirically, ecologists might need to use predictive methods  
240 to estimate the total number of interactions in a food web.

241 We used the flexible links model of MacDonald, Banville, and Poisot (2020) to predict the  
242 number of interactions from the number of species. The flexible links model, in contrast to other  
243 predictive models of the number of interactions, incorporates meaningful ecological constraints

244 into the prediction of  $L$ , namely the minimum  $S - 1$  and maximum  $S^2$  numbers of interactions  
 245 in food webs. It estimates the proportion of the  $S^2 - (S - 1)$  *flexible links* that are realized.  
 246 More precisely, this model states that the number of *realized* flexible links (or interactions)  
 247  $L_{FL}$  in a food web represents the number of realized interactions above the minimum (i.e.  $L =$   
 248  $L_{FL} + S - 1$ ) and is obtained from a beta-binomial distribution with  $S^2 - (S - 1)$  trials and  
 249 parameters  $\alpha = \mu e^\phi$  and  $\beta = (1 - \mu)e^\phi$ :

$$L_{FL} \sim \text{BB}(S^2 - (S - 1), \mu e^\phi, (1 - \mu)e^\phi), \quad (20)$$

250 where  $\mu$  is the average probability across food webs that a flexible link is realized, and  $\phi$  is the  
 251 concentration parameter around  $\mu$ .

252 We fitted the flexible links model on all food webs in our complete dataset, and estimated the  
 253 parameters of eq. 20 using a Hamiltonian Monte Carlo sampler with static trajectory (1 chain  
 254 and 3000 iterations):

$$[\mu, \phi | \mathbf{L}, \mathbf{S}] \propto \prod_{i=1}^m \text{BB}(L_i - (S_i - 1) | S_i^2 - (S_i - 1)), \mu e^\phi, (1 - \mu)e^\phi) \times \text{B}(\mu | 3, 7) \times \mathcal{N}(\phi | 3, 0.5), \quad (21)$$

255 where  $m$  is the number of food webs ( $m = 257$ ) and  $\mathbf{L}$  and  $\mathbf{S}$  are respectively the vectors of their  
 256 numbers of interactions and numbers of species. Our weakly-informative prior distributions  
 257 were chosen following MacDonald, Banville, and Poisot (2020), i.e. a beta distribution for  $\mu$   
 258 and a normal distribution for  $\phi$ . The Monte Carlo sampling of the posterior distribution was  
 259 conducted using the Julia library Turing v0.15.12.

260 The flexible links model is a generative model, i.e. it can generate plausible values of the pre-  
 261 dicted variable. We thus simulated 1000 values of  $L$  for different values of  $S$  using the joint  
 262 posterior distribution of our model parameters, and calculated the mean degree for each simu-  
 263 lated values. The resulting distributions are shown in the left panel of fig. 3 for three different  
 264 values of species richness. In the right panel of fig. 3, we show how the probability distribution  
 265 for the mean degree constraints can be used to generate a distribution of maximum entropy de-

266 gree distributions, since each simulated value of mean degree generates a different maximum  
267 entropy degree distribution.

268 [Figure 3 about here.]

## 269 **Heuristical models: Networks of maximum entropy**

### 270 **MaxEnt network models (types I and II)**

271 We define networks of maximum entropy as the configuration of the adjacency matrix with the  
272 highest SVD entropy under a set of constraints. As mentioned in Box 1, we used the SVD  
273 entropy as our measure of entropy since it has been shown to be a reliable measure of food-web  
274 complexity (Strydom, Dalla Riva, and Poisot 2021), in addition to having the required properties  
275 of a proper measure of information entropy. We thus seek to find the network with the highest  
276 complexity, or randomness, that exactly reproduces specified constraints on its structure. Our  
277 method is in contrast with maximum entropy graph models that predict a probability distribution  
278 on networks under soft or hard constraints (e.g., Park and Newman 2004; Cimini et al. 2019). We  
279 believe our approach to be more flexible, easier to compute, while allowing direct comparisons  
280 of empirical food webs with more complex networks with similar structure.

281 We built two types of MaxEnt network models: one based on connectance (type I MaxEnt net-  
282 work model) and the other based on the joint degree sequence (type II MaxEnt network model).  
283 They are based on the same constraints as the types I and II null models presented above. For  
284 both models, we used a simulated annealing algorithm with 4 chains, 2000 steps and an initial  
285 temperature of 0.2. For each chain, we first generated one random Boolean matrix with the same  
286 order (number of species) as empirical webs, while maintaining the total number of interactions  
287 (type I MaxEnt network model) or rows and columns sums (type II MaxEnt network model).  
288 These are our initial configurations. Then, we swapped interactions sequentially while main-  
289 taining the original connectance or the joint degree sequence for types I and II MaxEnt network  
290 model, respectively. Configurations with a higher SVD entropy than the previous one in the  
291 chain were always accepted, whereas they were accepted with a probability conditional to a de-

292 creasing temperature when lower. The final configuration with the highest SVD entropy among  
293 the four chains constitute our estimated MaxEnt network. Even though we decided to work with  
294 point estimates, it is possible to have a (non MaxEnt) probability distribution of networks when  
295 working with the entire chains after burn-in. For each network in our complete and abundance  
296 datasets, we estimated their configuration with maximum entropy using both types of MaxEnt  
297 network models.

## 298 Structure of MaxEnt networks

299 We compared different measures of the structure of empirical food webs with the ones of null,  
300 neutral, and MaxEnt networks [tbl. 1](#), and [2](#). For instance, we measured the nestedness  $\rho$  accord-  
301 ing to Staniczenko, Kopp, and Allesina (2013), i.e. using the spectral radius of the adjacency  
302 matrix. Nestedness indicates how much the diet of specialist species is a subset of the one of  
303 generalists, and is strongly associated with network modularity (Fortuna et al. 2010). We also  
304 measured the maximum trophic level *maxtl*, network diameter *diam* (i.e. the longest of the short-  
305 est paths between all species pairs; Albert and Barabasi 2002), the average maximum similarity  
306 between species pairs *MxSim*, the proportion of cannibal species *Cannib* (i.e. the proportion of  
307 self loops), and the proportion of omnivorous species *Omniv* (i.e. species whose preys are of  
308 different trophic levels). *MxSim*, *Cannib*, and *Omniv* are more deeply defined in Williams and  
309 Martinez (2000).

310 Overall, we found that models based on the joint degree sequence (type II null and MaxEnt  
311 network models) reproduced the structure of empirical networks much better than the ones based  
312 on connectance and the neutral model of relative abundances. On average, the type II MaxEnt  
313 network model was better at predicting nestedness ( $0.62 \pm 0.08$ ) than its corresponding null model  
314 ( $0.73 \pm 0.05$ ; empirical networks:  $0.63 \pm 0.09$ ), as well as the proportion of cannibal species.  
315 However, the type II null model was better at predicting network diameter and average maximum  
316 similarity between species pairs. Predictions were similar between both type II models for the  
317 maximum trophic level and the proportion of omnivorous species.

Table 1: Standardized mean differences of predicted network measures with empirical data for all networks ( $n = 257$ ). Positive (negative) values indicate that the measure is overestimated (underestimated) on average. Empirical networks are all food webs archived on Mangal, as well as the New Zealand and Tuesday lake food webs (our complete dataset). Null 1: Type I null model based on connectance. MaxEnt 1: Type I MaxEnt network model based on connectance. Null 2: Type II null model based on the joint degree sequence. MaxEnt 2: Type II MaxEnt network model based on the joint degree sequence.  $\rho$ : nestedness measured by the spectral radius of the adjacency matrix. *maxtl*: maximum trophic level. *diam*: network diameter. *MxSim*: average maximum similarity between species pairs. *Cannib*: proportion of cannibal species (self loops). *Omniv*: proportion of omnivorous species. *entropy*: SVD entropy.

model	rho	maxtl	diam	MxSim	Cannib	Omniv	entropy
null 1	-0.167	0.980	1.428	-0.502	2.007	1.493	0.056
MaxEnt 1	-0.226	0.831	1.274	-0.524	1.982	1.863	0.106
null 2	0.160	-0.125	0.016	0.007	1.078	0.559	-0.023
MaxEnt 2	-0.015	0.178	0.565	-0.282	0.698	0.589	0.058

Table 2: Standardized mean differences of predicted network measures with empirical data for networks having abundance data ( $n = 19$ ). Positive (negative) values indicate that the measure is overestimated (underestimated) on average. Empirical networks are all New Zealand and Tuesday lake food webs with abundance data (our abundance dataset). Null 1: Type I null model based on connectance. MaxEnt 1: Type I MaxEnt network model based on connectance. Null 2: Type II null model based on the joint degree sequence. MaxEnt 2: Type II MaxEnt network model based on the joint degree sequence.  $\rho$ : nestedness measured by the spectral radius of the adjacency matrix. *maxtl*: maximum trophic level. *diam*: network diameter. *MxSim*: average maximum similarity between species pairs. *Cannib*: proportion of cannibal species (self loops). *Omniv*: proportion of omnivorous species. *entropy*: SVD entropy.

model	rho	maxtl	diam	MxSim	Cannib	Omniv	entropy
neutral	0.367	-0.090	0.027	0.266	6.870	0.576	-0.083
null 1	-0.134	0.950	1.919	-0.369	2.077	0.614	0.068
MaxEnt 1	-0.229	1.020	1.946	-0.355	2.215	0.801	0.121
null 2	0.128	-0.115	-0.135	0.157	1.444	0.029	-0.021
MaxEnt 2	-0.010	0.054	0.243	-0.062	-0.038	0.083	0.038

318 The picture slightly changes when we consider another important property of ecological net-  
 319 works, i.e. their motifs profile. We measured the proportion of three-species motifs, which can

320 be considered as simple building blocks of ecological networks (Milo et al. 2002; Stouffer et al.  
321 2007). Motifs are the backbone of complex ecological networks from which network structure  
322 is built upon and play a crucial role in community dynamics and assembly (Bastolla et al. 2009;  
323 Stouffer and Bascompte 2011). In fig. 4, we show that the motifs profile of networks generated  
324 using the type II MaxEnt network model was very close to the one of empirical data, and that  
325 they made better predictions than the type II null model. This is also shown in fig. 5, where we  
326 see that relationships between motifs pairs of empirical food webs are very similar to the ones  
327 of the type II MaxEnt networks. Overall, these results suggest that the type II MaxEnt network  
328 model can reproduce many aspects of network structure. This highlights the importance of the  
329 joint degree sequence in shaping ecological networks.

330 [Figure 4 about here.]

331 [Figure 5 about here.]

332 Moreover, we found that empirical networks are close to their maximum entropy considering a  
333 fixed joint degree sequence (fig. S3). Empirical networks had an SVD entropy of  $0.89 \pm 0.04$ ,  
334 compared to type II MaxEnt networks which had values of  $0.94 \pm 0.03$ . The relationship between  
335 the SVD entropy of empirical food webs and their maximum entropy configuration is plotted in  
336 the last panel of fig. 6. We also found no relationship between the difference in SVD entropy  
337 between empirical and type II MaxEnt networks and species richness, the number of interactions,  
338 or connectance (fig. S4). Similarly, we found no correlation between the difference in SVD  
339 entropy and the Jaccard distance of empirical and type II MaxEnt networks.

340 Finally, in fig. 6 we also show how well empirical measures are predicted using the type II  
341 MaxEnt network model. In accordance with our previous results, we found that nestedness was  
342 very well predicted by our model. However, the type II MaxEnt network model overestimated  
343 the maximum trophic level and network diameter, especially for networks with high empirical  
344 values. In fig. S5, we show that the pairwise relationships between these four measures and  
345 species richness in empirical food webs are similar (in magnitude and sign) to the ones in type  
346 II MaxEnt networks.

348 **Conclusion**

349 The principle of maximum entropy is a robust mathematical method of finding least-biased  
 350 probability distributions that have some desired properties given by prior knowledge on a sys-  
 351 tem. We first used MaxEnt to predict the joint degree distribution and the degree distribution  
 352 of maximum entropy given known numbers of species and of interactions. We found that the  
 353 resulting joint degree distributions were more symmetric than the ones of empirical food webs,  
 354 which suggests that other constraints might be needed to improve those predictions. We also  
 355 used MaxEnt to predict networks of maximum entropy with a specified structure. These net-  
 356 works are the most complex, or random, given the constraints used. Likewise, we found that  
 357 knowledge of species richness and of the number of interactions were not sufficient to reproduce  
 358 many aspects of network structure found in nature. However, a model based on the entire joint  
 359 degree sequence, i.e. on the number of preys and predators for each species, gave more con-  
 360 vincing results. Our type II MaxEnt network model yielded better or similar predictions than  
 361 the type II null model, also based on the joint degree sequence, for most measures considered  
 362 including the motifs profile. This suggests that the joint degree sequence drives many aspects of  
 363 network structure. Indeed, considering our findings that empirical networks are close to max-  
 364 imum entropy for a given joint degree sequence, our results suggest that food-web topology is  
 365 configured almost entirely randomly around these marginal numbers of predators and preys.

366 Our method and results could be used for different purposes. First, they could be used as first-  
 367 order approximations of network structure when only state variables are known. This could  
 368 prove useful when predicting network structure at large spatial scales, where few ecological in-  
 369 formation is known at that scale. Second, they could be used as informative priors in Bayesian  
 370 analyses of the structure of ecological networks (e.g., Cirtwill et al. 2019). Third, they could  
 371 be used to make better predictions of pairwise species interactions by constraining the space of  
 372 feasible networks, as discussed in Strydom et al. (2021). Finally, they could be used as alter-  
 373 native null models of ecological networks to better understand ecological mechanisms driving

374 food-web structure. In that case, our model might need to be slightly adapted to give a proba-  
375 bility distribution of Boolean networks (in contrast with point estimates of maximum entropy  
376 networks).

377 One of the biggest challenges in using the principle of maximum entropy is to identify the set  
378 of state variables that best reproduce empirical data. We found that the numbers of preys and  
379 predators for each species are important state variables for the prediction of maximum entropy  
380 networks. However, our predictions overestimated some measures of network structure, espe-  
381 cially the maximum trophic level and network diameter. Therefore, we should continue playing  
382 the ecological detective to find these other topological constraints that would improve the pre-  
383 dictions of our MaxEnt network models.

## 384 Acknowledgments

385 We acknowledge that this study was conducted on land within the traditional unceded territory  
386 of the Saint Lawrence Iroquoian, Anishinabewaki, Mohawk, Huron-Wendat, and Omàmiwinini-  
387 wak nations. This work was supported by the Institute for Data Valorisation (IVADO) and the  
388 NSERC BIOS<sup>2</sup> CREATE program.

## 389 References

- 390 Albert, R., and A. L. Barabasi. 2002. “Statistical Mechanics of Complex Networks.” *Reviews*  
391 *of Modern Physics* 74 (1): 47–97. <https://doi.org/10.1103/RevModPhys.74.47>.
- 392 Banville, Francis, Steve Vissault, and Timothée Poisot. 2021. “Mangal.jl and EcologicalNet-  
393 works.jl: Two Complementary Packages for Analyzing Ecological Networks in Julia.” *Jour-*  
394 *nal of Open Source Software* 6 (61): 2721. <https://doi.org/10.21105/joss.02721>.
- 395 Bascompte, J., P. Jordano, C. J. Melian, and J. M. Olesen. 2003. “The Nested Assembly of  
396 Plant-Animal Mutualistic Networks.” *Proceedings of the National Academy of Sciences*  
397 *of the United States of America* 100 (16): 9383–87. <https://doi.org/10.1073/pnas.1633576100>.

- 399 Bastolla, Ugo, Miguel A. Fortuna, Alberto Pascual-Garcia, Antonio Ferrera, Bartolo Luque, and  
400 Jordi Bascompte. 2009. “The Architecture of Mutualistic Networks Minimizes Competition  
401 and Increases Biodiversity.” *Nature* 458 (7241): 1018–U91. <https://doi.org/10.1038/nature07950>.
- 403 Beck, Christian. 2009. “Generalised Information and Entropy Measures in Physics.” *Contemporary Physics* 50 (4): 495–510. <https://doi.org/10.1080/00107510902823517>.
- 405 Cimini, Giulio, Tiziano Squartini, Fabio Saracco, Diego Garlaschelli, Andrea Gabrielli, and  
406 Guido Caldarelli. 2019. “The Statistical Physics of Real-World Networks.” *Nature Reviews Physics* 1 (1): 58–71. <https://doi.org/10.1038/s42254-018-0002-6>.
- 408 Cirtwill, Alyssa R., Anna Eklof, Tomas Roslin, Kate Wootton, and Dominique Gravel. 2019.  
409 “A Quantitative Framework for Investigating the Reliability of Empirical Network Construc-  
410 tion.” *Methods in Ecology and Evolution* 10 (6): 902–11. <https://doi.org/10.1111/2041-210X.13180>.
- 412 Cohen, Joel E., Tomas Jonsson, and Stephen R. Carpenter. 2003. “Ecological Community  
413 Description Using the Food Web, Species Abundance, and Body Size.” *Proceedings of the National Academy of Sciences* 100 (4): 1781–86. <https://doi.org/10.1073/pnas.232715699>.
- 416 Delmas, Eva, Mathilde Besson, Marie Hélène Brice, Laura A. Burkle, Giulio V. Dalla Riva,  
417 Marie Josée Fortin, Dominique Gravel, et al. 2019. “Analysing Ecological Networks of  
418 Species Interactions.” *Biological Reviews*. <https://doi.org/10.1111/brv.12433>.
- 419 Dunning, Iain, Joey Huchette, and Miles Lubin. 2017. “JuMP: A Modeling Language for  
420 Mathematical Optimization.” *SIAM Review* 59 (2): 295–320. <https://doi.org/10.1137/15M1020575>.
- 422 Fortuna, M. A., and J. Bascompte. 2006. “Habitat Loss and the Structure of Plant-Animal  
423 Mutualistic Networks.” *Ecology Letters* 9 (3): 278–83. <https://doi.org/10.1111/j.1461-0248.2005.00868.x>.
- 425 Fortuna, M. A., Daniel B. Stouffer, Jens M. Olesen, Pedro Jordano, David Mouillot, Boris R.  
426 Krasnov, Robert Poulin, and Jordi Bascompte. 2010. “Nestedness Versus Modularity in

- 427 Ecological Networks: Two Sides of the Same Coin?” *Journal of Animal Ecology* 79 (4):  
428 811–17. <https://doi.org/10.1111/j.1365-2656.2010.01688.x>.
- 429 Frank, S. A., and E. Smith. 2011. “A Simple Derivation and Classification of Common Prob-  
430 ability Distributions Based on Information Symmetry and Measurement Scale.” *Journal of*  
431 *Evolutionary Biology* 24 (3): 469–84. <https://doi.org/10.1111/j.1420-9101.2010.02204.x>.
- 432 Harremoës, Peter, and Flemming Topsøe. 2001. “Maximum Entropy Fundamentals.” *Entropy*  
433 3 (3): 191–226. <https://doi.org/10.3390/e3030191>.
- 434 Harte, J., and Erica A. Newman. 2014. “Maximum Information Entropy: A Foundation for  
435 Ecological Theory.” *Trends in Ecology & Evolution* 29 (7): 384–89. <https://doi.org/10.1016/j.tree.2014.04.009>.
- 436 Harte, J., T. Zillio, E. Conlisk, and A. B. Smith. 2008. “Maximum Entropy and the State-  
437 Variable Approach to Macroecology.” *Ecology* 89 (10): 2700–2711. <https://doi.org/10.1890/07-1369.1>.
- 438 Hoorn, Pim van der, Gabor Lippner, and Dmitri Krioukov. 2018. “Sparse Maximum-Entropy  
439 Random Graphs with a Given Power-Law Degree Distribution.” *Journal of Statistical Physics*  
440 173 (3): 806–44. <https://doi.org/10.1007/s10955-017-1887-7>.
- 441 Hortal, Joaquín, Francesco de Bello, José Alexandre F. Diniz-Filho, Thomas M. Lewinsohn,  
442 Jorge M. Lobo, and Richard J. Ladle. 2015. “Seven Shortfalls That Beset Large-Scale  
443 Knowledge of Biodiversity.” *Annual Review of Ecology, Evolution, and Systematics* 46 (1):  
444 523–49. <https://doi.org/10.1146/annurev-ecolsys-112414-054400>.
- 445 Jaynes, E. T. 1957a. “Information Theory and Statistical Mechanics.” *Physical Review* 106 (4):  
446 620–30. <https://doi.org/10.1103/PhysRev.106.620>.
- 447 ———. 1957b. “Information Theory and Statistical Mechanics. II.” *Physical Review* 108 (2):  
448 171–90. <https://doi.org/10.1103/PhysRev.108.171>.
- 449 Khinchin, A. Ya. 2013. *Mathematical Foundations of Information Theory*. Courier Corpora-  
450 tion.

- 454 MacDonald, Arthur Andrew Meahan, Francis Banville, and Timothée Poisot. 2020. “Revisiting  
455 the Links-Species Scaling Relationship in Food Webs.” *Patterns* 0 (0). <https://doi.org/10.1016/j.patter.2020.100079>.
- 456
- 457 Martyushev, L. M., and V. D. Seleznev. 2006. “Maximum Entropy Production Principle in  
458 Physics, Chemistry and Biology.” *Physics Reports-Review Section of Physics Letters* 426  
459 (1): 1–45. <https://doi.org/10.1016/j.physrep.2005.12.001>.
- 460 Milo, R., S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. 2002. “Net-  
461 work Motifs: Simple Building Blocks of Complex Networks.” *Science* 298 (5594): 824–27.  
462 <https://doi.org/10.1126/science.298.5594.824>.
- 463 Park, Juyong, and M. E. J. Newman. 2004. “Statistical Mechanics of Networks.” *Physical*  
464 *Review E* 70 (6): 066117. <https://doi.org/10.1103/PhysRevE.70.066117>.
- 465 Phillips, Steven J., Robert P. Anderson, and Robert E. Schapire. 2006. “Maximum Entropy  
466 Modeling of Species Geographic Distributions.” *Ecological Modelling* 190 (3): 231–59.  
467 <https://doi.org/10.1016/j.ecolmodel.2005.03.026>.
- 468 Pielou, Evelyn C. 1975. “Ecological Diversity.” In. 574.524018 P5.
- 469 Poisot, Timothee, Benjamin Baiser, Jennifer A. Dunne, Sonia Kefi, Francois Massol, Nico-  
470 las Mouquet, Tamara N. Romanuk, Daniel B. Stouffer, Spencer A. Wood, and Dominique  
471 Gravel. 2016. “Mangal - Making Ecological Network Analysis Simple.” *Ecography* 39 (4):  
472 384–90. <https://doi.org/10.1111/ecog.00976>.
- 473 Pomeranz, Justin Page, Ross M. Thompson, Timothée Poisot, Jon S. Harding, and Justin P. F.  
474 Pomeranz. 2018. “Data from: Inferring Predator-Prey Interactions in Food Webs.” Dryad.  
475 <https://doi.org/10.5061/DRYAD.K59M37F>.
- 476 Shannon, C. E. 1948. “A Mathematical Theory of Communication.” *The Bell System Technical*  
477 *Journal* 27 (3): 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
- 478 Staniczenko, Phillip P. A., Jason C. Kopp, and Stefano Allesina. 2013. “The Ghost of Nested-  
479 ness in Ecological Networks.” *Nature Communications* 4 (1): 1391. <https://doi.org/10.1038/ncomms2422>.
- 480

- 481 Stock, Michiel, Timothée Poisot, and Bernard De Baets. 2021. “Optimal Transportation Theory  
482 for Species Interaction Networks.” *Ecology and Evolution* 11 (9): 3841–55. <https://doi.org/10.1002/ece3.7254>.
- 483
- 484 Stouffer, Daniel B., and Jordi Bascompte. 2011. “Compartmentalization Increases Food-Web  
485 Persistence.” *Proceedings of the National Academy of Sciences* 108 (9): 3648–52. <https://doi.org/10.1073/pnas.1014353108>.
- 486
- 487 Stouffer, Daniel B., Juan Camacho, Wenxin Jiang, and Luis A. Nunes Amaral. 2007. “Evidence  
488 for the Existence of a Robust Pattern of Prey Selection in Food Webs.” *Proceedings of the  
489 Royal Society B-Biological Sciences* 274 (1621): 1931–40. <https://doi.org/10.1098/rspb.2007.0571>.
- 490
- 491 Strydom, Tanya, Michael D. Catchen, Francis Banville, Dominique Caron, Gabriel Dansereau,  
492 Philippe Desjardins-Proulx, Norma R. Forero-Muñoz, et al. 2021. “A Roadmap Towards  
493 Predicting Species Interaction Networks (across Space and Time).” *Philosophical Trans-  
494 actions of the Royal Society B: Biological Sciences* 376 (1837): 20210063. <https://doi.org/10.1098/rstb.2021.0063>.
- 495
- 496 Strydom, Tanya, Giulio V. Dalla Riva, and Timothee Poisot. 2021. “SVD Entropy Reveals the  
497 High Complexity of Ecological Networks.” *Frontiers in Ecology and Evolution* 9: 623141.  
498 <https://doi.org/10.3389/fevo.2021.623141>.
- 499
- 500 Williams, Richard J. 2011. “Biology, Methodology or Chance? The Degree Distributions of  
501 Bipartite Ecological Networks.” *PLOS ONE* 6 (3): e17645. <https://doi.org/10.1371/journal.pone.0017645>.
- 502
- 503 Williams, Richard J., and Neo D. Martinez. 2000. “Simple Rules Yield Complex Food Webs.”  
*Nature* 404 (6774): 180–83. <https://doi.org/10.1038/35004572>.

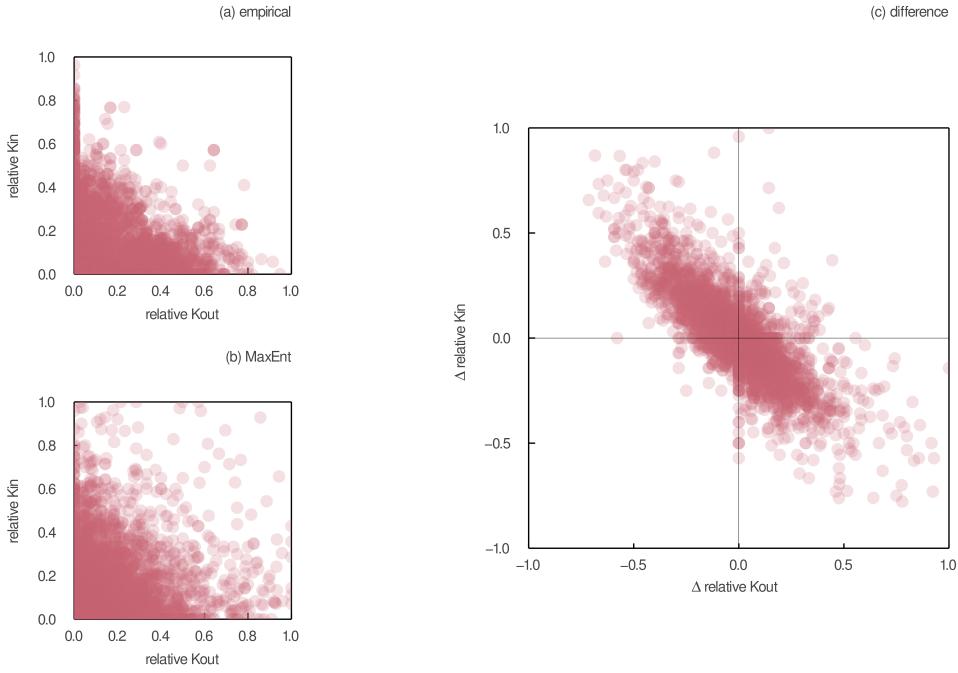


Figure 1: Relative number of predators ( $k_{in}$ ) as a function of the relative number of preys ( $k_{out}$ ) for each species in (a) empirical and (b) maximum entropy joint degree sequences. Empirical networks include all food webs archived on Mangal, as well as the New-Zealand and Tuesday lake datasets (our complete dataset). The predicted joint degree sequences were obtained after sampling one realization of the joint degree distribution of maximum entropy for each network. (c) Difference between predicted and empirical values when species are ordered according to their total degrees. In all panels, each dot corresponds to a single species in one of the network of our complete dataset.

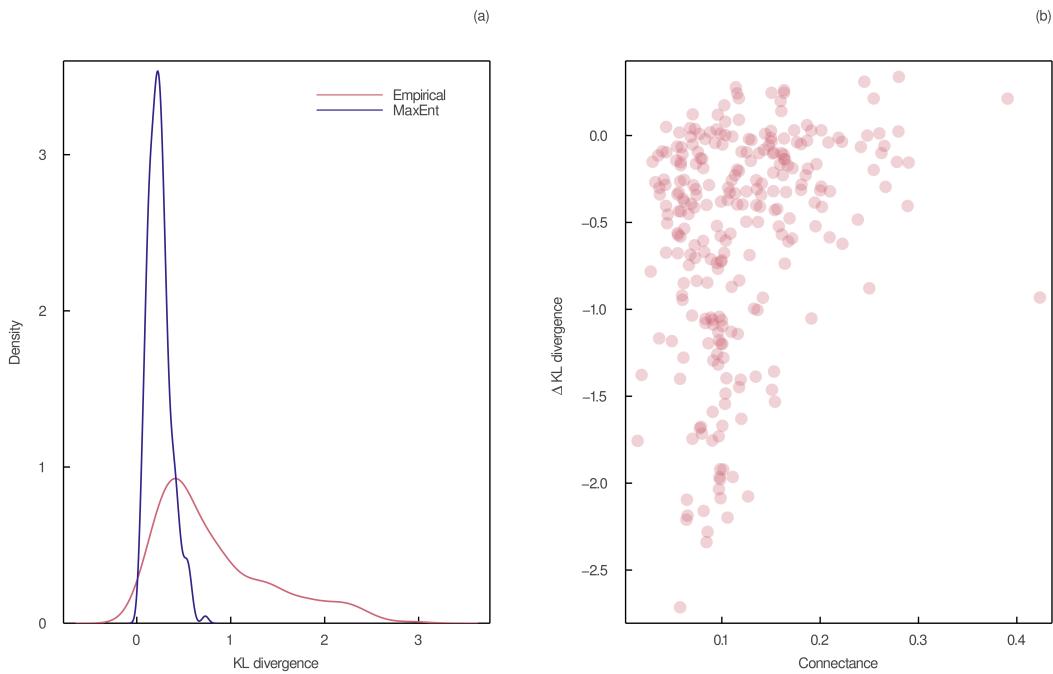


Figure 2: (a) Probability density of KL divergence between in and out-degree sequences of empirical and maximum entropy joint degree sequences. (b) Difference between the KL divergence of empirical and predicted networks as a function of connectance. In both panels, empirical networks include all food webs archived on Mangal, as well as the New-Zealand and Tuesday lake datasets (our complete dataset). The joint degree sequence of simulated networks was obtained after sampling one realization of the joint degree distribution of maximum entropy for each network.

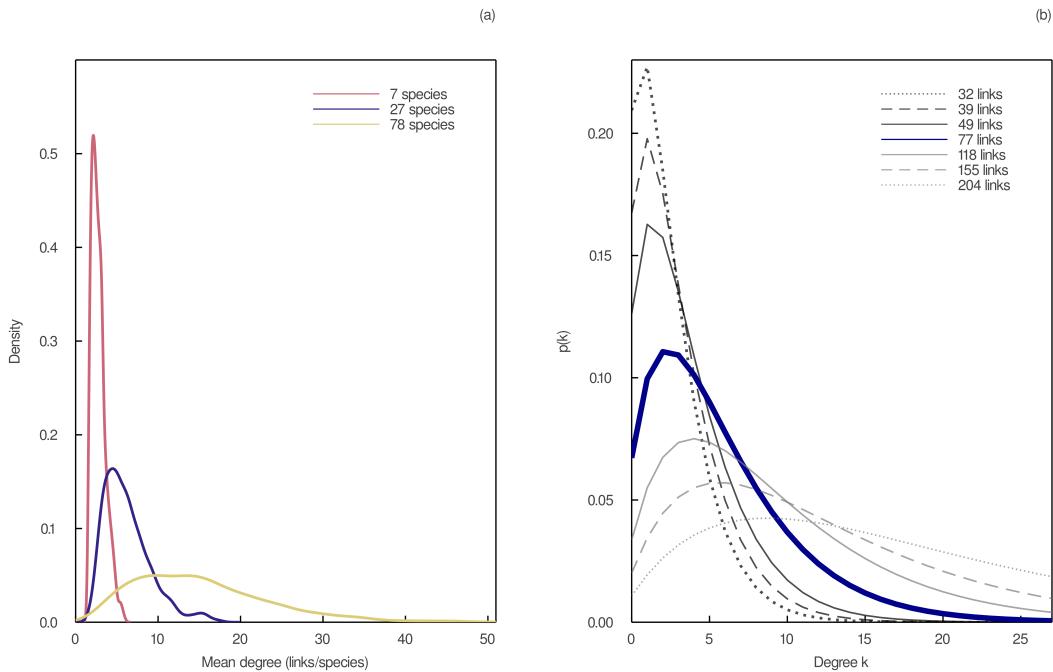


Figure 3: (a) Probability density of the predicted mean degree for different values of species richness. The number of interactions was predicted using the flexible links model fitted to all empirical networks in our complete dataset. (b) Degree distributions of maximum entropy for a network of 27 species and different numbers of interactions. The numbers of interactions correspond to the lower and upper bounds of the 67, 89, and 97 percentile intervals (PI), as well as the median, of the counterfactuals of the flexible links model.

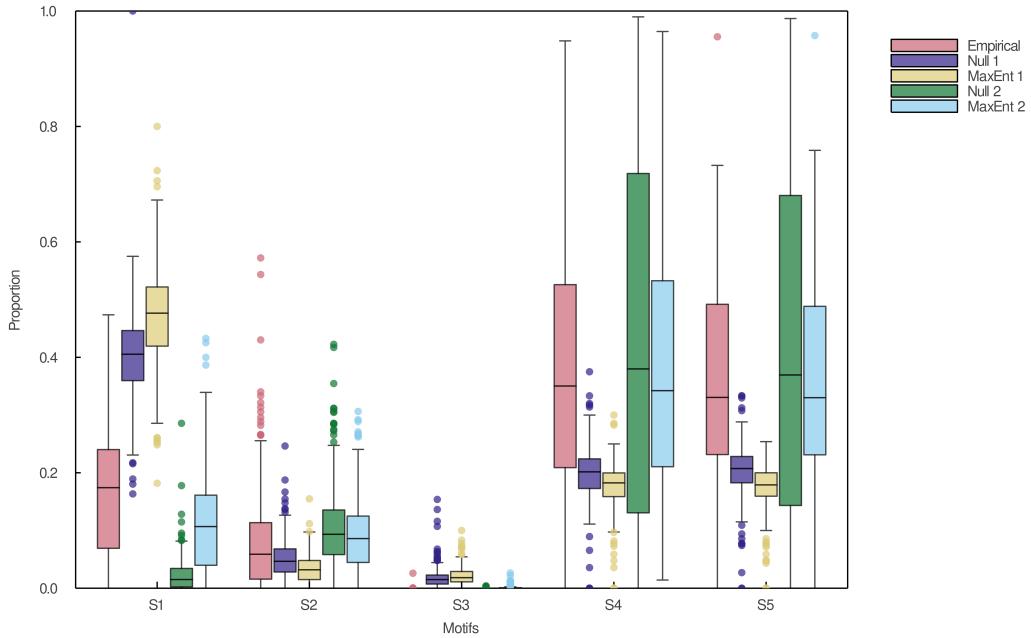


Figure 4: Motifs profile of empirical, maximum entropy, and null food webs. Empirical networks include all food webs archived on Mangal, as well as the New-Zealand and Tuesday lake datasets (our complete dataset). Maximum entropy networks were estimated using a simulated annealing algorithm to find the network of maximum SVD entropy while maintaining the connectance (type I MaxEnt network model) or the joint degree sequence (type II MaxEnt network model). The predictions of the types I and II null models are also plotted. Boxplots display the median proportions of each motif (middle horizontal lines), as well as the first (bottom horizontal lines) and third (top horizontal lines) quartiles. Vertical lines encompass all data points that fall within 1.5 times the interquartile range from both quartiles, and dots are data points that fall outside this range. Motifs names are from Stouffer et al. (2007).

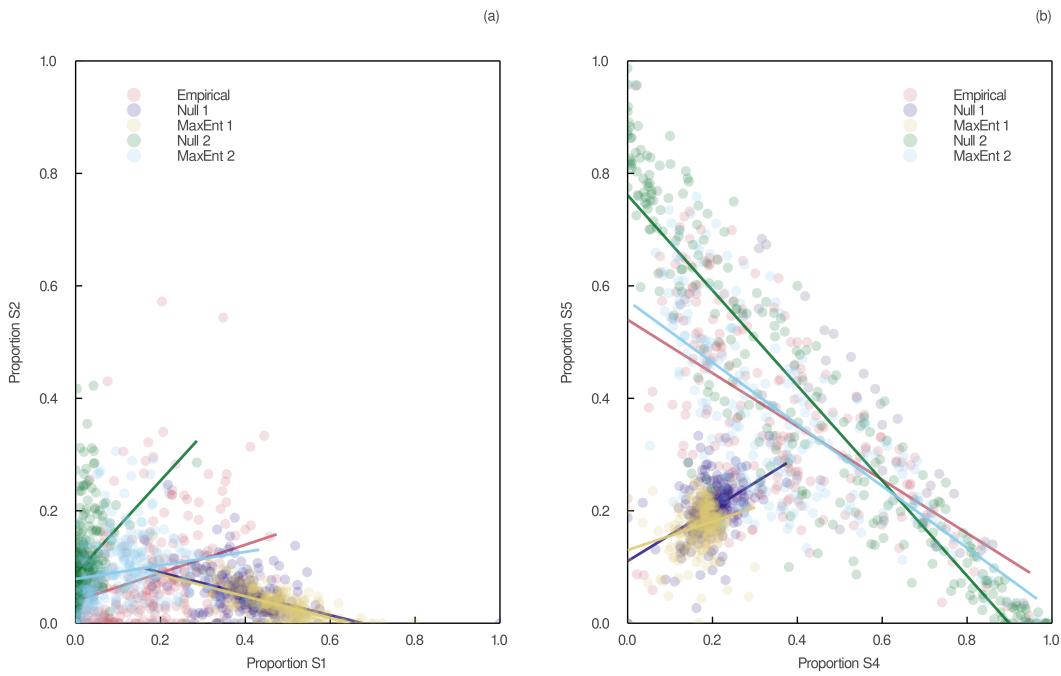


Figure 5: Relationship between motifs proportions of empirical, maximum entropy, and null food webs. Empirical networks include all food webs archived on Mangal, as well as the New-Zealand and Tuesday lake datasets (our complete dataset). Maximum entropy networks were estimated using a simulated annealing algorithm to find the network of maximum SVD entropy while maintaining the connectance (type I MaxEnt network model) or the joint degree sequence (type II MaxEnt network model). The predictions of the types I and II null models are also plotted. Regression lines are plotted in each panel. Motifs names are from Stouffer et al. (2007).

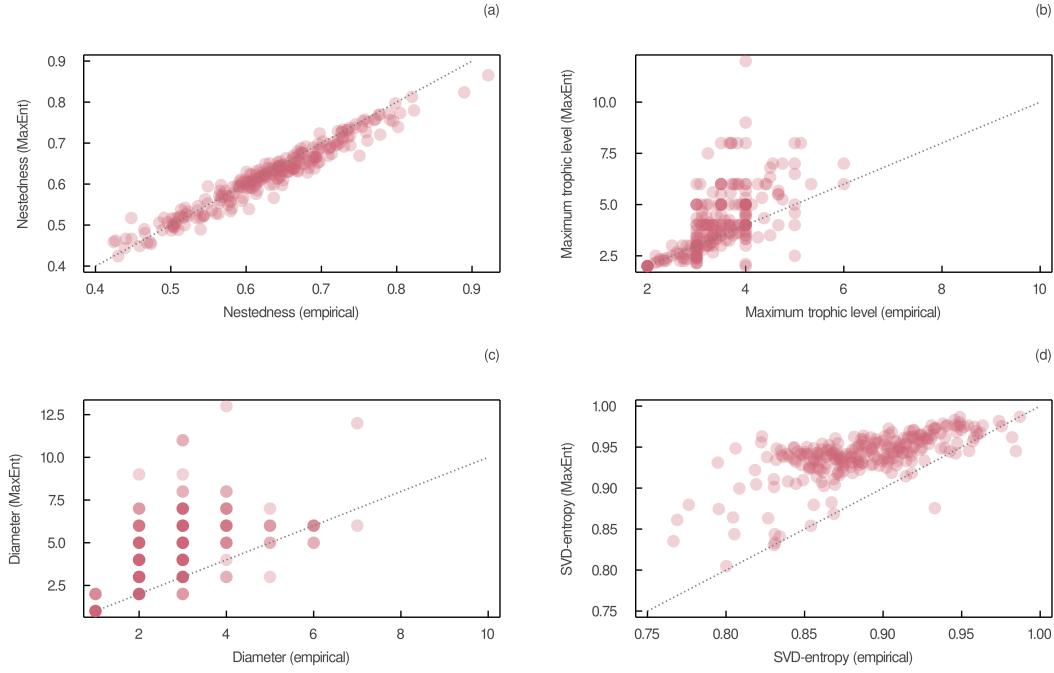


Figure 6: Relationship between the structure of empirical and maximum entropy food webs. Empirical networks include all food webs archived on Mangal, as well as the New-Zealand and Tuesday lake datasets (our complete dataset). Maximum entropy networks were estimated using a simulated annealing algorithm to find the network of maximum SVD entropy while maintaining the joint degree sequence (type II MaxEnt network model). (a) Nestedness (estimated using the spectral radius of the adjacency matrix), (b) the maximum trophic level, (c) the network diameter, and (d) the SVD entropy were measured on these empirical and maximum entropy food webs. The identity line is plotted in each panel.