

What constrains food webs? A maximum entropy framework for predicting their structure with minimal biases

Francis Banville^{1,2,3} Dominique Gravel^{2,3} Timothée Poisot^{1,3}

¹ Université de Montréal ² Université de Sherbrooke ³ Quebec Centre for Biodiversity Science

Correspondance to:

Francis Banville — francis.banville@umontreal.ca

Abstract: Food webs are complex ecological networks whose structure is both ecologically and statistically constrained, with many network properties being correlated with each other. Despite the recognition of these invariable relationships in food webs, the use of the principle of maximum entropy (MaxEnt) in network ecology is still rare. This is surprising considering that MaxEnt is a renowned and rigorous statistical tool precisely designed for understanding and predicting many different types of constrained systems. Precisely, this principle asserts that the least-biased probability distribution of a system's property, constrained by prior knowledge about that system, is the one with maximum information entropy. MaxEnt has been proven useful in many ecological modelling problems, but its application in food webs and other ecological networks is limited. Here we show how MaxEnt can be used to derive many food-web properties both analytically and heuristically. First, we show how the joint degree distribution (the joint probability distribution of the numbers of prey and predators for each species in the network) can be derived analytically using the number of species and the number of interactions in food webs. Second, we present a heuristic and flexible approach of finding a network's adjacency matrix (the network's representation in matrix format) based on simulated annealing and SVD entropy. We built two heuristic models using the connectance and the joint degree sequence as statistical constraints, respectively. We compared both models' predictions against corresponding null and neutral models commonly used in network ecology using open access data of terrestrial and aquatic food webs sampled globally ($N = 257$). We found that the heuristic model constrained by the joint degree sequence was a good predictor of many measures of food-web structure, especially the nestedness and motifs distribution. Specifically, our results suggest that the structure of terrestrial and aquatic food webs is mainly driven by their joint degree distribution.

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
17 be used for both of these purposes, i.e. to make predictions of network structure and to better
18 understand processes shaping ecological networks. Food-web properties that can be derived
19 using MaxEnt are varied and pertain to different elements of the network (i.e. at the species
20 (node), the interaction (edge) or the community (network) levels). For example, at the species
21 level, MaxEnt can be used to predict the distribution of trophic levels among species, as well as
22 the distribution of species' vulnerability (number of predators) and generality (number of preys).
23 By contrast, at the interaction level, predictions can be made on the distribution of interaction
24 strengths in weighted food webs. At the community level, it can give us a probability distribution
25 of potential networks or of specific measures of their emerging structure. The applicability and
26 performance of this method depend on the ecological information available and on our capacity
27 to find the set of state variables that best represent natural systems and to translate them into
28 appropriate statistical constraints. MaxEnt has been used in a wide range of disciplines, from
29 thermodynamics to chemistry and biology (Martyushev and Seleznev 2006). It has also been

30 proven useful in ecology, e.g. in species distribution models (Phillips, Anderson, and Schapire
31 2006) and macroecological models (Harte et al. 2008; Harte and Newman 2014). Maximizing a
32 measure of entropy ensures that the derived probability distributions are unique and least biased
33 under the set of constraints used. These constraints are built using state variables, i.e. variables
34 that represent the macrostate of the system. Having a validated maximum entropy model for the
35 system at hand allows us to make rigorous predictions using a minimal amount of data, as well
36 as helping us describe the most important factors driving that system.

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

49 In this contribution, we used two complementary approaches to predict the structure of food
50 webs using the principle of maximum entropy. We then compared our predictions against em-
51 pirical data and null and neutral models commonly used in network ecology. The first approach
52 consists in deriving constrained probability distributions of given network properties directly.
53 We derived the joint degree distribution (a probability distribution) of maximum entropy using
54 only the number of species S and the number of interactions L as state variables. Then, we
55 predicted the degree distribution of maximum entropy directly from the joint degree distribu-
56 tion since the first is the sum of the marginal distributions of the second (a species' degree is
57 the sum of its in and out-degrees). Because of the scarcity of empirical data on the number of
58 interactions in ecological networks, we present a method to predict L from S (Box 1), thus al-

59 lowing the prediction of the joint degree distribution from S solely. In turn, the second approach
60 consists in finding, using different constraints, the adjacency matrix of maximum entropy from
61 which network properties can be measured. To do so, we used a flexible and heuristic approach
62 based on simulated annealing to find networks *close* to maximum entropy. As discussed above,
63 our choice of algorithm stands from the very nature of food webs (i.e. simple directed networks
64 allowing self-loops) that makes the analytical derivation of a maximum entropy graph model
65 difficult. We first built our type I MaxEnt network model constrained by the connectance of
66 the network (i.e. the ratio L/S^2). A comparison of this model against empirical data indicated
67 that connectance alone was not sufficient to predict many aspects of network structure. For this
68 reason, we built our type II MaxEnt network model, which instead uses the whole joint degree
69 sequence as a constraint. Overall, we found that this second model was much better at predicting
70 food-web structure than the one based on connectance.

71 **The principle of maximum entropy: A primer for ecologists**

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

84 Entropy measures the amount of information given by the outcome of a random variable. Many
85 measures of entropy have been developed in physics (Beck 2009), but only a fraction of them
86 could be used as an optimization measure with the principle of maximum entropy. According

87 to Beck (2009) and Khinchin (2013), a measure of entropy H should satisfy four properties
 88 in the discrete case: (1) it should be a function of a probability distribution $p(n)$ only; (2) it
 89 should be maximized when $p(n)$ is uniform; (3) it should not be influenced by outcomes with a
 90 null probability; and (4) it should be independent of the order of information acquisition. The
 91 Shannon's entropy (Shannon 1948)

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

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

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

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

103 Shannon's entropy can only be calculated on conventional probability distributions such as the
 104 joint degree distribution. This is an issue when working with ecological networks, which are
 105 frequently represented as adjacency matrices. For this reason, in this contribution, we also use
 106 the SVD entropy as a measure of entropy, which is an application of Shannon's entropy to the
 107 relative non-zero singular values of a truncated singular value decomposition (t-SVD; Strydom,
 108 Dalla Riva, and Poisot 2021) of a food web's Boolean adjacency matrix. We measured SVD
 109 entropy as follows:

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

110 where s_i are the relative singular values of the adjacency matrix ($s_i = \sigma_i / \sum_{i=1}^R \sigma_i$, where σ_i
 111 are the singular values). Note that the distribution of relative singular values is analogous to a
 112 probability distribution, with $0 < s_i < 1$ and $\sum s_i = 1$. Thus, this measure also satisfies all
 113 four properties above-mentioned, while being a proper measure of the internal complexity of
 114 food webs (Strydom, Dalla Riva, and Poisot 2021). Following Strydom, Dalla Riva, and Poisot
 115 (2021), we standardized this measure with the rank R of the matrix (i.e. $J / \ln(R)$) to account
 116 for the difference in dimensions between networks (*sensu* Pielou's evenness; Pielou 1975). We
 117 will show how SVD entropy can be used to predict the network of maximum entropy (i.e. of
 118 maximum complexity) heuristically given a predetermined set of constraints.

119 Data and code

120 We tested our MaxEnt models (both approaches) against open food-web data queried from three
 121 different sources and integrated into what we call our *complete dataset*. First, all food webs
 122 archived on `mangal.io` (Poisot et al. 2016; Banville, Vissault, and Poisot 2021) were directly
 123 queried from the database ($n = 235$). Most ecological networks archived on Mangal are multi-
 124 layer networks, i.e. networks that describe different types of interactions. We kept all networks
 125 whose interactions were mainly of predation and herbivory types, and removed the largest net-
 126 work ($S = 714$) for computational efficiency reasons. Then, to this set we added food webs
 127 from two different sources: the New-Zealand dataset ($n = 21$; Pomeranz et al. 2018) and the
 128 Tuesday lake dataset ($n = 2$; Cohen, Jonsson, and Carpenter 2003). Of these two datasets, 19
 129 networks had data on species' relative abundances that were used in the neutral model presented
 130 in a following subsection. These networks are part of what we call our *abundance dataset*, which
 131 is a subset of our complete dataset.

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

¹³⁴ **Analytical models: Measures of maximum entropy**

¹³⁵ **Joint degree distribution**

¹³⁶ The joint degree distribution $p(k_{in}, k_{out})$ is a joint discrete probability distribution describing
¹³⁷ the probability that a species has k_{in} predators and k_{out} preys, with k_{in} and $k_{out} \in [0, S]$. Basal
¹³⁸ species (e.g., plants) have a k_{out} of 0, whereas top predators have a k_{in} of 0. In contrast, the
¹³⁹ maximum number of preys and predators a species can have is set by the number of species S
¹⁴⁰ in the food web. Here we show how the joint degree distribution of maximum entropy can be
¹⁴¹ obtained given knowledge of S and L .

¹⁴² We want to maximize Shannon's entropy

$$H = - \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S p(k_{in}, k_{out}) \log p(k_{in}, k_{out}) \quad (4)$$

¹⁴³ subject to the following constraints:

$$g_1 = \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S p(k_{in}, k_{out}) = 1; \quad (5)$$

$$g_2 = \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S k_{in} p(k_{in}, k_{out}) = \langle k_{in} \rangle = \frac{L}{S}; \quad (6)$$

$$g_3 = \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S k_{out} p(k_{in}, k_{out}) = \langle k_{out} \rangle = \frac{L}{S}. \quad (7)$$

¹⁴⁴ The first constraint g_1 is our normalizing constraint, whereas the other two (g_2 and g_3) fix the
¹⁴⁵ average of the marginal distributions of k_{in} and k_{out} to the linkage density L/S . It is important
¹⁴⁶ to notice that $\langle k_{in} \rangle = \langle k_{out} \rangle$ because every edge is associated to a predator and a prey. Therefore,
¹⁴⁷ without any further constraints, we expect the joint degree distribution of maximum entropy to
¹⁴⁸ be a symmetric probability distribution with regards to k_{in} and k_{out} . However, this does not mean
¹⁴⁹ that the joint degree *sequence* will be symmetric, since the joint degree sequence is essentially
¹⁵⁰ a random realization of its probabilistic counterpart.

₁₅₁ The joint probability distribution of maximum entropy given these constraints is found using the
₁₅₂ method of Lagrange multipliers. To do so, we seek to maximize the following expression:

$$F = H - \lambda_1(g_1 - 1) - \lambda_2\left(g_2 - \frac{L}{S}\right) - \lambda_3\left(g_3 - \frac{L}{S}\right), \quad (8)$$

₁₅₃ where λ_1 , λ_2 , and λ_3 are the Lagrange multipliers. The probability distribution that maximizes
₁₅₄ entropy is obtained by finding these values. As pointed out in the introduction, F is just Shan-
₁₅₅ non's entropy to which we added terms that each sums to zero (our constraints). F is maximized
₁₅₆ by setting to 0 its partial derivative with respect to $p(k_{in}, k_{out})$. Because the derivative of a con-
₁₅₇ stant is zero, this gives us:

$$\frac{\partial H}{\partial p(k_{in}, k_{out})} = \lambda_1 \frac{\partial g_1}{\partial p(k_{in}, k_{out})} + \lambda_2 \frac{\partial g_2}{\partial p(k_{in}, k_{out})} + \lambda_3 \frac{\partial g_3}{\partial p(k_{in}, k_{out})}. \quad (9)$$

₁₅₈ Evaluating the partial derivatives with respect to $p(k_{in}, k_{out})$, we obtain:

$$-\log p(k_{in}, k_{out}) - 1 = \lambda_1 + \lambda_2 k_{in} + \lambda_3 k_{out}. \quad (10)$$

₁₅₉ Then, solving eq. 10 for $p(k_{in}, k_{out})$, we obtain:

$$p(k_{in}, k_{out}) = \frac{e^{-\lambda_2 k_{in} - \lambda_3 k_{out}}}{Z}, \quad (11)$$

₁₆₀ where $Z = e^{1+\lambda_1}$ is called the partition function. The partition function ensures that probabilities
₁₆₁ sum to 1 (our normalization constraint). It can be expressed in terms of λ_2 and λ_3 as follows:

$$Z = \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S e^{-\lambda_2 k_{in} - \lambda_3 k_{out}}. \quad (12)$$

₁₆₂ After substituting $p(k_{in}, k_{out})$ in eq. 6 and eq. 7, we get a nonlinear system of two equations and
₁₆₃ two unknowns:

$$\frac{1}{Z} \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S k_{in} e^{-\lambda_2 k_{in} - \lambda_3 k_{out}} = \frac{L}{S}; \quad (13)$$

$$\frac{1}{Z} \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S k_{out} e^{-\lambda_2 k_{in} - \lambda_3 k_{out}} = \frac{L}{S}. \quad (14)$$

¹⁶⁴ We solved eq. 13 and eq. 14 numerically using the Julia library JuMP.jl v0.21.8 (Dunning,
¹⁶⁵ Huchette, and Lubin 2017). JuMP.jl supports nonlinear optimization problems by providing
¹⁶⁶ exact second derivatives that increase the accuracy and performance of its solvers. The estimated
¹⁶⁷ values of λ_2 and λ_3 can be substituted in eq. 11 to have a more workable expression for the joint
¹⁶⁸ degree distribution.

¹⁶⁹ We predicted the joint degree distribution of maximum entropy for each food web in our com-
¹⁷⁰ plete dataset, i.e. using their numbers of species and numbers of interactions as state variables.
¹⁷¹ We then sampled one realization of the degree sequence for each network using the probabilities
¹⁷² given by the joint degree distribution. In fig. 1 (left panels), we show the relationship between
¹⁷³ k_{out} and k_{in} standardized by the number of species in their networks, for empirical and maximum
¹⁷⁴ entropy joint degree distributions. We see that our model predicts higher values of generality
¹⁷⁵ and vulnerability compared to empirical food webs (i.e. relative values of k_{out} and k_{in} closer to
¹⁷⁶ 1). However, plotting the difference between predicted and empirical values for each species
¹⁷⁷ gives a different perspective. The right panel of fig. 1 presents these differences when species
¹⁷⁸ are ordered by their total degree in their network (i.e. by the sum of their in and out-degrees).
¹⁷⁹ Indeed, our predicted joint degree sequences have the same number of species as their empiri-
¹⁸⁰ cal counterparts, but they are species agnostic; in other words, instead of predicting a value for
¹⁸¹ each species directly, we predicted the entire joint degree sequence without taking into account
¹⁸² species' identity. When we associate predictions and empirical data according to their rank in
¹⁸³ total degrees, we see that species predicted to have a higher generality (number of preys)
¹⁸⁴ generally have a lower vulnerability (number of preys) than what is observed (and conversely).
¹⁸⁵ In fig. S1, we show how these differences change when species are instead ordered by their
¹⁸⁶ out-degrees (left panel) and in-degrees (right panel), respectively.

187

[Figure 1 about here.]

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

194

[Figure 2 about here.]

195 Degree distribution

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

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

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

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

$$p(k) = \frac{e^{-\lambda_2 k}}{Z}, \quad (15)$$

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

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

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

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

$$\frac{1}{Z} \sum_{k=0}^S k e^{-\lambda_2 k} = \frac{2L}{S} \quad (16)$$

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

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

225 **Box 1 - Working with predicted numbers of interactions**

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

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

232 into the prediction of L , namely the minimum $S - 1$ and maximum S^2 numbers of interactions
 233 in food webs. It estimates the proportion of the $S^2 - (S - 1)$ *flexible links* that are realized.
 234 More precisely, this model states that the number of *realized* flexible links (or interactions)
 235 L_{FL} in a food web represents the number of realized interactions above the minimum (i.e. $L =$
 236 $L_{FL} + S - 1$) and is obtained from a beta-binomial distribution with $S^2 - (S - 1)$ trials and
 237 parameters $\alpha = \mu e^\phi$ and $\beta = (1 - \mu) e^\phi$:

$$L_{FL} \sim \text{BB}(S^2 - (S - 1), \mu e^\phi, (1 - \mu) e^\phi), \quad (17)$$

238 where μ is the average probability across food webs that a flexible link is realized, and ϕ is the
 239 concentration parameter around μ .

240 We fitted the flexible links model on all food webs in our complete dataset, and estimated the
 241 parameters of eq. 17 using a Hamiltonian Monte Carlo sampler with static trajectory (1 chain
 242 and 3000 iterations):

$$[\mu, \phi | \mathbf{L}, \mathbf{S}] \propto \prod_{i=1}^m \text{BB}(L_i - (S_i - 1) | S_i^2 - (S_i - 1)), \mu e^\phi, (1 - \mu) e^\phi) \times \text{B}(\mu | 3, 7) \times \mathcal{N}(\phi | 3, 0.5), \quad (18)$$

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

248 The flexible links model is a generative model, i.e. it can generate plausible values of the pre-
 249 dicted variable. We thus simulated 1000 values of L for different values of S using the joint
 250 posterior distribution of our model parameters (eq. 18), and calculated the mean degree for
 251 each simulated values. The resulting distributions are shown in the left panel of fig. 3 for three
 252 different values of species richness. In the right panel of fig. 3, we show how the probability
 253 distribution for the mean degree constraints can be used to generate a distribution of maximum

254 entropy degree distributions, since each simulated value of mean degree generates a different
255 maximum entropy degree distribution (eq. 15 and eq. 16).

256 [Figure 3 about here.]

257 **Box 2 - Corresponding null and neutral models**

258 **Null models (types I and II)**

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

$$p(i \rightarrow j) = \frac{L}{S^2}. \quad (19)$$

262 The second is the type II null model of Bascompte et al. (2003), in which the probability of
263 interaction is instead given by

$$p(i \rightarrow j) = \frac{1}{2} \left(\frac{k_{in}(j)}{S} + \frac{k_{out}(i)}{S} \right), \quad (20)$$

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

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

274 **Neutral model**

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

$$p(i \rightarrow j) \propto \frac{n_i}{N} \times \frac{n_j}{N}, \quad (21)$$

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

281 **Heuristical models: Networks of maximum entropy**

282 **MaxEnt network models (types I and II)**

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

293 We built two types of MaxEnt network models: one based on connectance (type I MaxEnt net-
294 work model) and the other based on the joint degree sequence (type II MaxEnt network model).
295 They are based on the same constraints as the types I and II null models presented above. For
296 both models, we used a simulated annealing algorithm with 4 chains, 2000 steps and an initial

temperature of 0.2. For each chain, we first generated one random Boolean matrix with the same order (number of species) as empirical webs, while maintaining the total number of interactions (type I MaxEnt network model) or rows and columns sums (type II MaxEnt network model). These are our initial configurations. Then, we swapped interactions sequentially while maintaining the original connectance or the joint degree sequence for types I and II MaxEnt network model, respectively. Configurations with a higher SVD entropy than the previous one in the chain were always accepted, whereas they were accepted with a probability conditional to a decreasing temperature when lower. The final configuration with the highest SVD entropy among the four chains constitute our estimated MaxEnt network. Even though we decided to work with point estimates, it is possible to have a (non MaxEnt) probability distribution of networks when working with the entire chains after burn-in. For each network in our complete and abundance datasets, we estimated their configuration with maximum entropy using both types of MaxEnt network models.

Structure of MaxEnt networks

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

Overall, we found that models based on the joint degree sequence (type II null and MaxEnt network models) reproduced the structure of empirical networks much better than the ones based on connectance and the neutral model of relative abundances. On average, the type II MaxEnt

325 network model was better at predicting nestedness (0.62 ± 0.08) than its corresponding null model
 326 (0.73 ± 0.05 ; empirical networks: 0.63 ± 0.09), as well as the proportion of cannibal species.
 327 However, the type II null model was better at predicting network diameter and average maximum
 328 similarity between species pairs. Predictions were similar between both type II models for the
 329 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. ρ : 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. ρ : 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

model	rho	maxtl	diam	MxSim	Cannib	Omniv	entropy
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

330 The picture slightly changes when we consider another important property of ecological net-
 331 works, i.e. their motifs profile. We measured the proportion of three-species motifs, which can
 332 be considered as simple building blocks of ecological networks (Milo et al. 2002; Stouffer et al.
 333 2007). Motifs are the backbone of complex ecological networks from which network structure
 334 is built upon and play a crucial role in community dynamics and assembly (Bastolla et al. 2009;
 335 Stouffer and Bascompte 2011). In fig. 4, we show that the motifs profile of networks generated
 336 using the type II MaxEnt network model was very close to the one of empirical data, and that
 337 they made better predictions than the type II null model. This is also shown in fig. 5, where we
 338 see that relationships between motifs pairs of empirical food webs are very similar to the ones
 339 of the type II MaxEnt networks. Overall, these results suggest that the type II MaxEnt network
 340 model can reproduce many aspects of network structure. This highlights the importance of the
 341 joint degree sequence in shaping ecological networks.

342 [Figure 4 about here.]

343 [Figure 5 about here.]

344 Moreover, we found that empirical networks are close to their maximum entropy considering a
 345 fixed joint degree sequence (fig. S3). Empirical networks had an SVD entropy of 0.89 ± 0.04 ,
 346 compared to type II MaxEnt networks which had values of 0.94 ± 0.03 . The relationship between
 347 the SVD entropy of empirical food webs and their maximum entropy configuration is plotted in
 348 the last panel of fig. 6. We also found no relationship between the difference in SVD entropy
 349 between empirical and type II MaxEnt networks and species richness, the number of interactions,
 350 or connectance (fig. S4). Similarly, we found no correlation between the difference in SVD
 351 entropy and the Jaccard distance of empirical and type II MaxEnt networks.

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

359 [Figure 6 about here.]

360 Conclusion

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

378 Our method and results could be used for different purposes. First, they could be used as first-

379 order approximations of network structure when only state variables are known. This could
380 prove useful when predicting network structure at large spatial scales, where few ecological in-
381 formation is known at that scale. Second, they could be used as informative priors in Bayesian
382 analyses of the structure of ecological networks (e.g., Cirtwill et al. 2019). Third, they could
383 be used to make better predictions of pairwise species interactions by constraining the space of
384 feasible networks, as discussed in Strydom et al. (2021). Finally, they could be used as alter-
385 native null models of ecological networks to better understand ecological mechanisms driving
386 food-web structure. In that case, our model might need to be slightly adapted to give a proba-
387 bility distribution of Boolean networks (in contrast with point estimates of maximum entropy
388 networks).

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

396 Acknowledgments

397 We acknowledge that this study was conducted on land within the traditional unceded territory
398 of the Saint Lawrence Iroquoian, Anishinabewaki, Mohawk, Huron-Wendat, and Omàmiwinini-
399 wak nations. This work was supported by the Institute for Data Valorisation (IVADO) and the
400 Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Re-
401 search and Training Experience (CREATE) program, through the Computational Biodiversity
402 Science and Services (BIOS²) program.

403 **References**

- 404 Albert, R., and A. L. Barabasi. 2002. “Statistical Mechanics of Complex Networks.” *Reviews*
405 *of Modern Physics* 74 (1): 47–97. <https://doi.org/10.1103/RevModPhys.74.47>.
- 406 Banville, Francis, Steve Vissault, and Timothée Poisot. 2021. “Mangal.jl and EcologicalNet-
407 works.jl: Two Complementary Packages for Analyzing Ecological Networks in Julia.” *Jour-*
408 *nal of Open Source Software* 6 (61): 2721. <https://doi.org/10.21105/joss.02721>.
- 409 Bascompte, J., P. Jordano, C. J. Melian, and J. M. Olesen. 2003. “The Nested Assembly of
410 Plant-Animal Mutualistic Networks.” *Proceedings of the National Academy of Sciences*
411 *of the United States of America* 100 (16): 9383–87. [https://doi.org/10.1073/pnas.](https://doi.org/10.1073/pnas.1633576100)
412 [1633576100](https://doi.org/10.1073/pnas.1633576100).
- 413 Bastolla, Ugo, Miguel A. Fortuna, Alberto Pascual-Garcia, Antonio Ferrera, Bartolo Luque, and
414 Jordi Bascompte. 2009. “The Architecture of Mutualistic Networks Minimizes Competition
415 and Increases Biodiversity.” *Nature* 458 (7241): 1018–U91. <https://doi.org/10.1038/nature07950>.
- 417 Beck, Christian. 2009. “Generalised Information and Entropy Measures in Physics.” *Contem-*
418 *porary Physics* 50 (4): 495–510. <https://doi.org/10.1080/00107510902823517>.
- 419 Cimini, Giulio, Tiziano Squartini, Fabio Saracco, Diego Garlaschelli, Andrea Gabrielli, and
420 Guido Caldarelli. 2019. “The Statistical Physics of Real-World Networks.” *Nature Reviews*
421 *Physics* 1 (1): 58–71. <https://doi.org/10.1038/s42254-018-0002-6>.
- 422 Cirtwill, Alyssa R., Anna Eklof, Tomas Roslin, Kate Wootton, and Dominique Gravel. 2019.
423 “A Quantitative Framework for Investigating the Reliability of Empirical Network Construc-
424 tion.” *Methods in Ecology and Evolution* 10 (6): 902–11. <https://doi.org/10.1111/2041-210X.13180>.
- 426 Cohen, Joel E., Tomas Jonsson, and Stephen R. Carpenter. 2003. “Ecological Community
427 Description Using the Food Web, Species Abundance, and Body Size.” *Proceedings of*
428 *the National Academy of Sciences* 100 (4): 1781–86. <https://doi.org/10.1073/pnas.232715699>.

- 430 Delmas, Eva, Mathilde Besson, Marie Hélène Brice, Laura A. Burkle, Giulio V. Dalla Riva,
431 Marie Josée Fortin, Dominique Gravel, et al. 2019. “Analysing Ecological Networks of
432 Species Interactions.” *Biological Reviews*. <https://doi.org/10.1111/brv.12433>.
- 433 Dunning, Iain, Joey Huchette, and Miles Lubin. 2017. “JuMP: A Modeling Language for
434 Mathematical Optimization.” *SIAM Review* 59 (2): 295–320. <https://doi.org/10.1137/15M1020575>.
- 436 Fortuna, M. A., and J. Bascompte. 2006. “Habitat Loss and the Structure of Plant-Animal
437 Mutualistic Networks.” *Ecology Letters* 9 (3): 278–83. <https://doi.org/10.1111/j.1461-0248.2005.00868.x>.
- 439 Fortuna, M. A., Daniel B. Stouffer, Jens M. Olesen, Pedro Jordano, David Mouillot, Boris R.
440 Krasnov, Robert Poulin, and Jordi Bascompte. 2010. “Nestedness Versus Modularity in
441 Ecological Networks: Two Sides of the Same Coin?” *Journal of Animal Ecology* 79 (4):
442 811–17. <https://doi.org/10.1111/j.1365-2656.2010.01688.x>.
- 443 Frank, S. A., and E. Smith. 2011. “A Simple Derivation and Classification of Common Prob-
444 ability Distributions Based on Information Symmetry and Measurement Scale.” *Journal of*
445 *Evolutionary Biology* 24 (3): 469–84. <https://doi.org/10.1111/j.1420-9101.2010.02204.x>.
- 447 Harremoës, Peter, and Flemming Topsøe. 2001. “Maximum Entropy Fundamentals.” *Entropy*
448 3 (3): 191–226. <https://doi.org/10.3390/e3030191>.
- 449 Harte, J., and Erica A. Newman. 2014. “Maximum Information Entropy: A Foundation for
450 Ecological Theory.” *Trends in Ecology & Evolution* 29 (7): 384–89. <https://doi.org/10.1016/j.tree.2014.04.009>.
- 452 Harte, J., T. Zillio, E. Conlisk, and A. B. Smith. 2008. “Maximum Entropy and the State-
453 Variable Approach to Macroecology.” *Ecology* 89 (10): 2700–2711. <https://doi.org/10.1890/07-1369.1>.
- 455 Hoorn, Pim van der, Gabor Lippner, and Dmitri Krioukov. 2018. “Sparse Maximum-Entropy
456 Random Graphs with a Given Power-Law Degree Distribution.” *Journal of Statistical Physics*
457 173 (3): 806–44. <https://doi.org/10.1007/s10955-017-1887-7>.

- 458 Hortal, Joaquín, Francesco de Bello, José Alexandre F. Diniz-Filho, Thomas M. Lewinsohn,
459 Jorge M. Lobo, and Richard J. Ladle. 2015. “Seven Shortfalls That Beset Large-Scale
460 Knowledge of Biodiversity.” *Annual Review of Ecology, Evolution, and Systematics* 46 (1):
461 523–49. <https://doi.org/10.1146/annurev-ecolsys-112414-054400>.
- 462 Jaynes, E. T. 1957a. “Information Theory and Statistical Mechanics.” *Physical Review* 106 (4):
463 620–30. <https://doi.org/10.1103/PhysRev.106.620>.
- 464 ———. 1957b. “Information Theory and Statistical Mechanics. II.” *Physical Review* 108 (2):
465 171–90. <https://doi.org/10.1103/PhysRev.108.171>.
- 466 Khinchin, A. Ya. 2013. *Mathematical Foundations of Information Theory*. Courier Corpora-
467 tion.
- 468 MacDonald, Arthur Andrew Meahan, Francis Banville, and Timothée Poisot. 2020. “Revisiting
469 the Links-Species Scaling Relationship in Food Webs.” *Patterns* 0 (0). <https://doi.org/10.1016/j.patter.2020.100079>.
- 471 Martyushev, L. M., and V. D. Seleznev. 2006. “Maximum Entropy Production Principle in
472 Physics, Chemistry and Biology.” *Physics Reports-Review Section of Physics Letters* 426
473 (1): 1–45. <https://doi.org/10.1016/j.physrep.2005.12.001>.
- 474 Milo, R., S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. 2002. “Net-
475 work Motifs: Simple Building Blocks of Complex Networks.” *Science* 298 (5594): 824–27.
476 <https://doi.org/10.1126/science.298.5594.824>.
- 477 Park, Juyong, and M. E. J. Newman. 2004. “Statistical Mechanics of Networks.” *Physical
478 Review E* 70 (6): 066117. <https://doi.org/10.1103/PhysRevE.70.066117>.
- 479 Phillips, Steven J., Robert P. Anderson, and Robert E. Schapire. 2006. “Maximum Entropy
480 Modeling of Species Geographic Distributions.” *Ecological Modelling* 190 (3): 231–59.
481 <https://doi.org/10.1016/j.ecolmodel.2005.03.026>.
- 482 Pielou, Evelyn C. 1975. “Ecological Diversity.” In. 574.524018 P5.
- 483 Poisot, Timothee, Benjamin Baiser, Jennifer A. Dunne, Sonia Kefi, Francois Massol, Nico-
484 las Mouquet, Tamara N. Romanuk, Daniel B. Stouffer, Spencer A. Wood, and Dominique

- 485 Gravel. 2016. “Mangal - Making Ecological Network Analysis Simple.” *Ecography* 39 (4):
486 384–90. <https://doi.org/10.1111/ecog.00976>.
- 487 Pomeranz, Justin Page, Ross M. Thompson, Timothée Poisot, Jon S. Harding, and Justin P. F.
488 Pomeranz. 2018. “Data from: Inferring Predator-Prey Interactions in Food Webs.” Dryad.
489 <https://doi.org/10.5061/DRYAD.K59M37F>.
- 490 Shannon, C. E. 1948. “A Mathematical Theory of Communication.” *The Bell System Technical
491 Journal* 27 (3): 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
- 492 Staniczenko, Phillip P. A., Jason C. Kopp, and Stefano Allesina. 2013. “The Ghost of Nested-
493 ness in Ecological Networks.” *Nature Communications* 4 (1): 1391. <https://doi.org/10.1038/ncomms2422>.
- 495 Stock, Michiel, Timothée Poisot, and Bernard De Baets. 2021. “Optimal Transportation Theory
496 for Species Interaction Networks.” *Ecology and Evolution* 11 (9): 3841–55. <https://doi.org/10.1002/ece3.7254>.
- 498 Stouffer, Daniel B., and Jordi Bascompte. 2011. “Compartmentalization Increases Food-Web
499 Persistence.” *Proceedings of the National Academy of Sciences* 108 (9): 3648–52. <https://doi.org/10.1073/pnas.1014353108>.
- 501 Stouffer, Daniel B., Juan Camacho, Wenxin Jiang, and Luis A. Nunes Amaral. 2007. “Evidence
502 for the Existence of a Robust Pattern of Prey Selection in Food Webs.” *Proceedings of the
503 Royal Society B-Biological Sciences* 274 (1621): 1931–40. <https://doi.org/10.1098/rspb.2007.0571>.
- 505 Strydom, Tanya, Michael D. Catchen, Francis Banville, Dominique Caron, Gabriel Dansereau,
506 Philippe Desjardins-Proulx, Norma R. Forero-Muñoz, et al. 2021. “A Roadmap Towards
507 Predicting Species Interaction Networks (across Space and Time).” *Philosophical Transactions of the Royal Society B: Biological Sciences* 376 (1837): 20210063. <https://doi.org/10.1098/rstb.2021.0063>.
- 510 Strydom, Tanya, Giulio V. Dalla Riva, and Timothee Poisot. 2021. “SVD Entropy Reveals the
511 High Complexity of Ecological Networks.” *Frontiers in Ecology and Evolution* 9: 623141.
512 <https://doi.org/10.3389/fevo.2021.623141>.

513 Williams, Richard J. 2011. "Biology, Methodology or Chance? The Degree Distributions of
514 Bipartite Ecological Networks." *PLOS ONE* 6 (3): e17645. <https://doi.org/10.1371/journal.pone.0017645>.

515
516 Williams, Richard J., and Neo D. Martinez. 2000. "Simple Rules Yield Complex Food Webs."
517 *Nature* 404 (6774): 180–83. <https://doi.org/10.1038/35004572>.

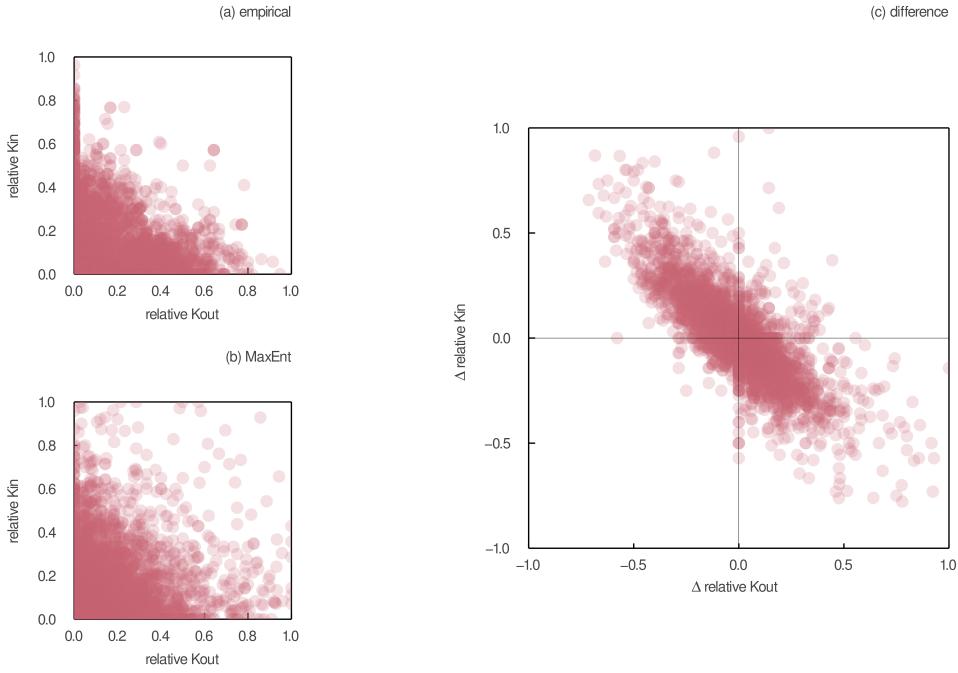


Figure 1: Relative number of predators (k_{in}) as a function of the relative number of preys (k_{out}) for each species in (a) empirical and (b) maximum entropy joint degree sequences. Empirical networks include all food webs archived on Mangal, as well as the New-Zealand and Tuesday lake datasets (our 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.

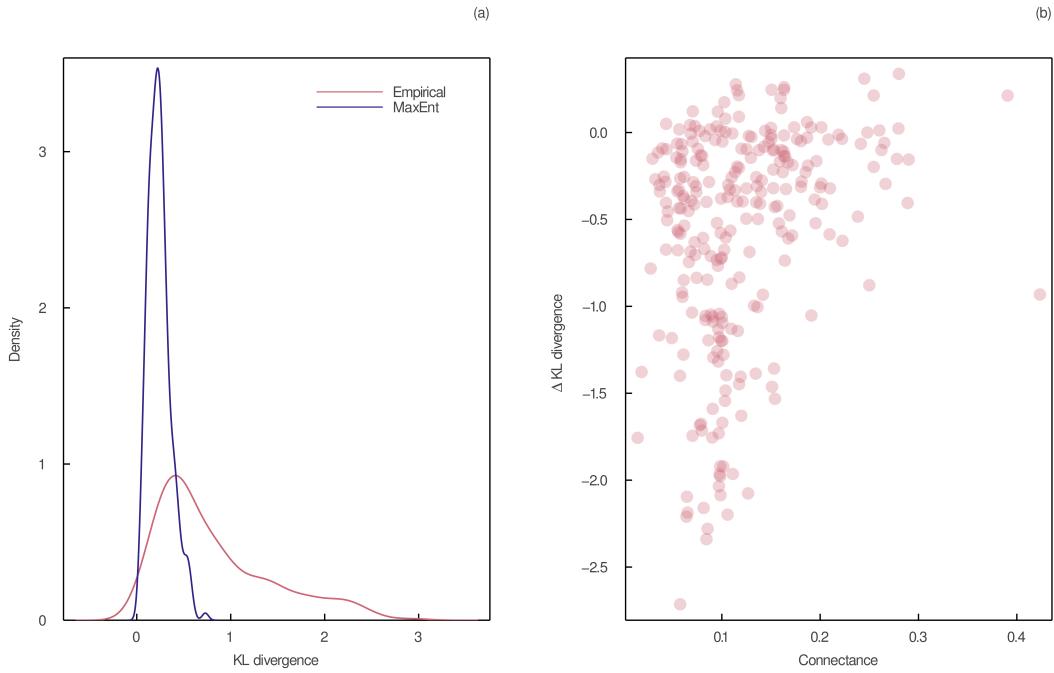


Figure 2: (a) Probability density of KL divergence between in and out-degree sequences of empirical and maximum entropy joint degree sequences. (b) Difference between the KL divergence of empirical and predicted networks as a function of connectance. In both panels, empirical networks include all food 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.

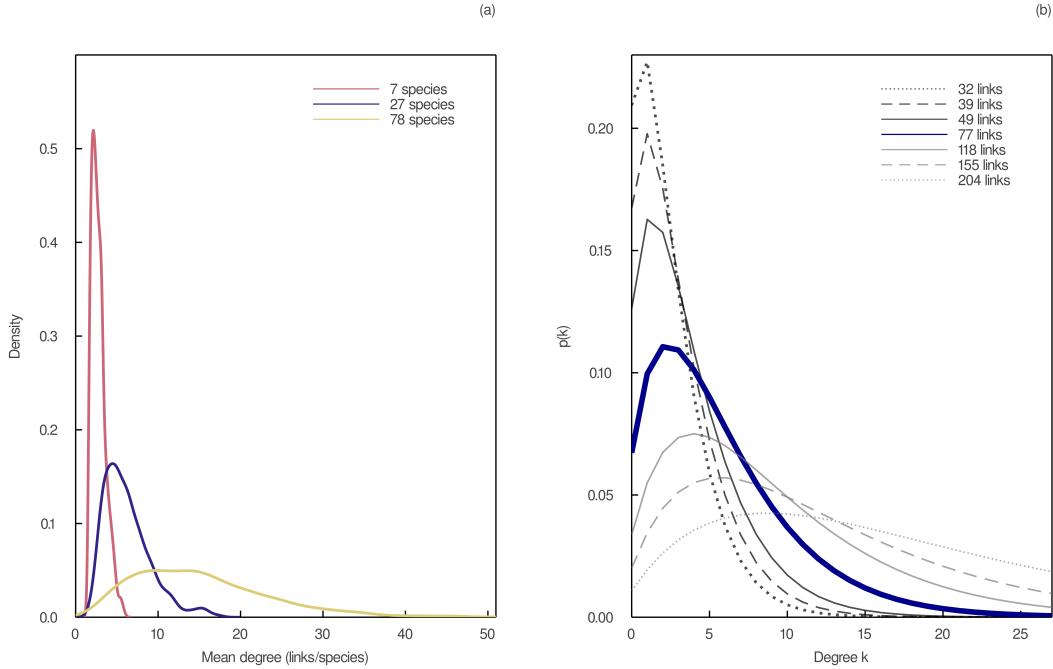


Figure 3: (a) Probability density of the predicted mean degree for different values of species richness. The number of interactions was predicted using the flexible links model fitted to all empirical networks in our complete dataset (eq. 18). (b) Degree distributions of maximum entropy for a network of $S = 27$ species and different numbers of interactions L . 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 (in blue), of the counterfactuals of the flexible links model. Each degree distribution was obtained using eq. 15 after solving numerically eq. 16 using different values for the mean degree constraint $2L/S$.

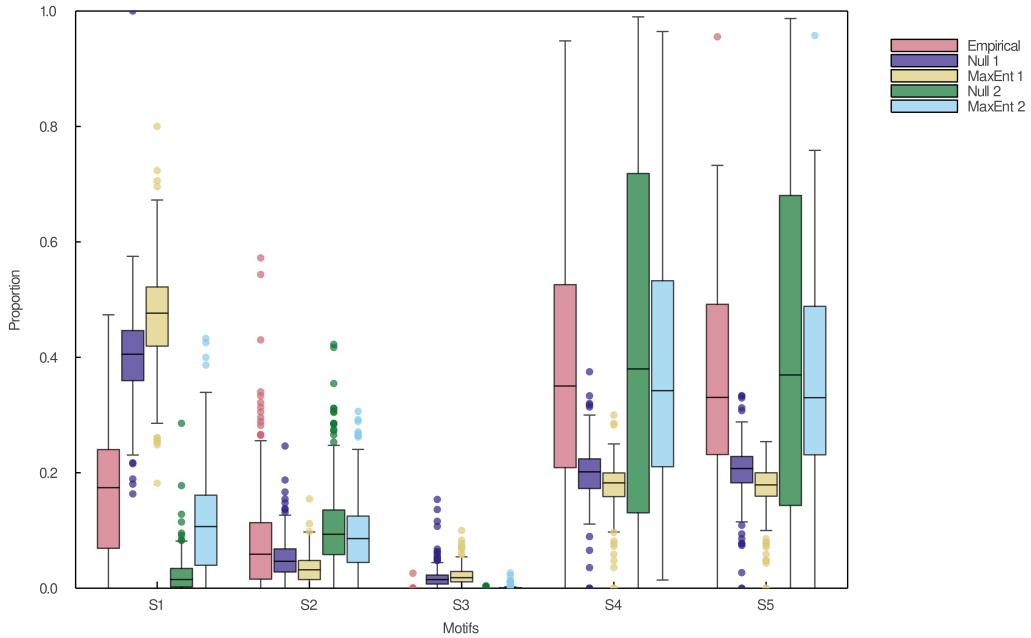


Figure 4: Motifs profile of empirical, maximum entropy, and null food webs. Empirical networks include all food webs archived on Mangal, as well as the New-Zealand and Tuesday lake datasets (our complete dataset). Maximum entropy networks were estimated using a simulated annealing algorithm to find the 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).

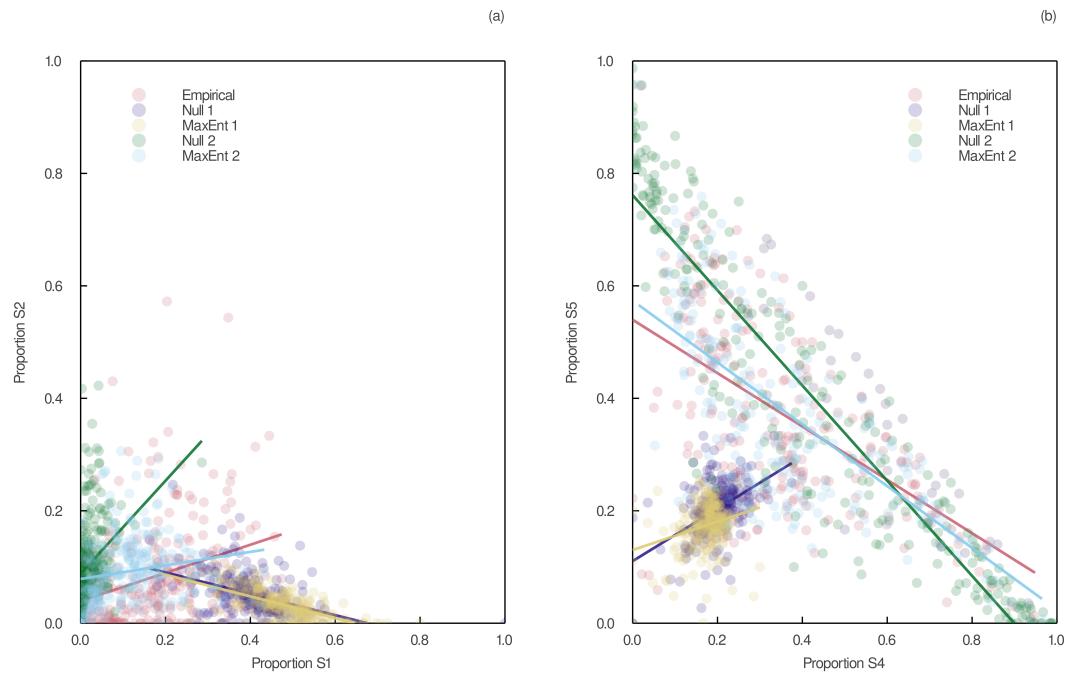


Figure 5: Relationship between motifs proportions of empirical, maximum entropy, and null food webs. Empirical networks include all food webs archived on Mangal, as well as the New-Zealand and Tuesday lake datasets (our complete dataset). Maximum entropy networks were estimated using a simulated annealing 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).

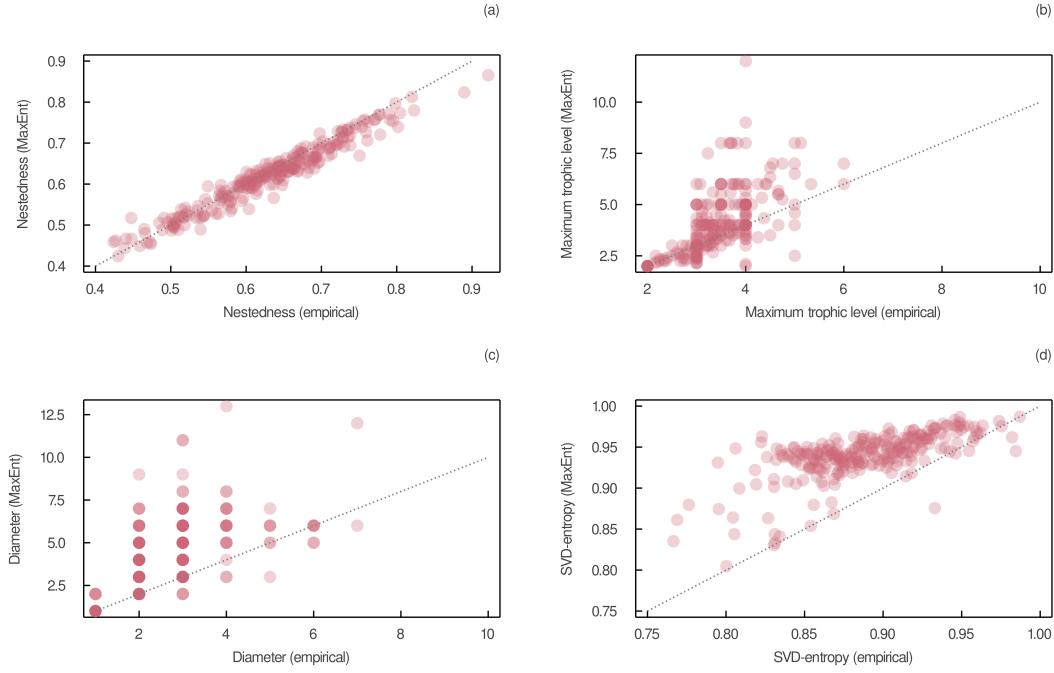


Figure 6: Relationship between the structure of empirical and maximum entropy food webs. Empirical networks include all food webs archived on Mangal, as well as the New-Zealand and Tuesday lake datasets (our complete dataset). Maximum entropy networks were estimated using a simulated annealing algorithm 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.