

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

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

<sup>1</sup> Département de sciences biologiques, Université de Montréal, Montreal, Quebec, Canada    <sup>2</sup> Département de biologie, Université de Sherbrooke, Sherbrooke, Quebec, Canada    <sup>3</sup> Quebec Centre for Biodiversity Science, Quebec, Canada

## 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 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.

**Author Summary:** Predator and prey species form complex networks of energy flow that drive the functioning and dynamics of ecosystems. Because of their high number and variability, sampling and documenting these feeding relationships are particularly challenging, especially at large spatial and temporal scales. Fortunately, many properties of food webs can be studied without knowing every trophic interaction in a network by using suitable computational methods. This is notably the case of their emerging structure, which could be driven by just a few ecological and biological variables despite the large number of mechanisms shaping ecological networks. In this contribution, we present a novel framework to identify these variables by applying the principle of maximum entropy to the analysis of food-web structure. We show that the number of prey and predators for each species in a food web is a fundamental property that shape many aspects of the network, while also being predictable by our model and available data. Our approach, which can be applied to a wide range of biological and ecological networks, can be used to make better predictions of food webs across space, in addition to contributing to our understanding of the ecological processes shaping predatory interactions at the community level.

# **1 Introduction**

## **2 The constrained structure of ecological networks**

3 A variety of measures of the structure of ecological networks have been used to describe the organization of  
4 species interactions in a biological community (Delmas *et al.* (2019)). These measures provide valuable  
5 information on the functioning of ecosystems and their responses to environmental change (e.g. Pascual &  
6 Dunne (2006), Gómez *et al.* (2011)). For instance, two main properties of a network are its order (the number  
7 of nodes or species  $S$ ) and its size (the number of edges or species interactions  $L$ ), whose ratio  $L/S^2$  (the  
8 proportion of possible interactions that are realized) is called connectance. In food webs, Dunne *et al.* (2002)  
9 showed that a high connectance can promote the robustness of the community to species lost. Moreover, a  
10 different investigation of network structure revealed that plant–pollinator and seed-disperser networks have a  
11 highly nested structure that can promote species persistence (Bascompte *et al.* (2003)). Despite the recognition  
12 of the ecological significance of network structure, our knowledge of the properties of ecological networks  
13 remains limited and is unevenly distributed across space (Poisot *et al.* (2021)). This can largely be attributed to  
14 the difficulties encountered when sampling interactions (Jordano (2016)), which impede our ability to describe  
15 network structure on a global scale. Moreover, these complex systems (Williams & Martinez (2000)) have an  
16 emerging structure with intricate dynamics (i.e. non-linear behaviors that cannot be predicted from individual  
17 interactions). Their multi-faceted nature and high level of complexity represent a challenge when studying the  
18 ecological processes driving their structure, especially given the strong relationship between many of their  
19 properties. For example, nestedness and modularity are highly correlated in ecological networks (Fortuna *et al.*  
20 (2010)) and many emerging properties are linked to connectance (Poisot & Gravel (2014)). In light of these  
21 observations, it is difficult to assess whether the presumed effects of a particular measure on network structure  
22 and behavior are the artifacts of other, perhaps simpler, measures.

23 One way to address this issue is to recognize that food webs and other ecological networks are constrained  
24 systems. In other words, the space of possible network configurations is shaped by our partial knowledge of  
25 their structure. For example, there is a finite set of networks with a given number of species and interactions  
26 (i.e. a given connectance). As shown by Poisot & Gravel (2014), connectance can constrain many aspects of  
27 network structure such as the degree distribution (i.e. the probability distribution of the number of interactions  
28 realized by a species). Because of the scaling relationship between the total number of species and interactions,  
29 network structure is also largely constrained by its order (food webs with high species richness typically have a

30 low connectance compared to smaller networks, MacDonald *et al.* (2020)). Other measures, such as the  
31 maximum trophic level (i.e. the maximum number of times energy is transformed along food chains through  
32 biomass consumption), may also constrain the space of feasible networks, e.g. by imposing a limit on energy  
33 transfer (Williams & Martinez (2004)).

34 Prior knowledge of the structure of ecological networks, even partial, can be useful in the current context of data  
35 scarcity about species interactions. The Eltonian shortfall, which describes the gap between our present-day  
36 knowledge of ecological networks and a comprehensive understanding of interactions (Hortal *et al.* (2015)), can  
37 be partially alleviated using known information about a network's attributes. First, knowing a given set of  
38 network properties (e.g. the number of species and interactions), it is possible to identify various structures that  
39 meet these constraints and estimate their respective probabilities of occurrence. For instance, a set of adjacency  
40 matrices (i.e. a representation of interactions in matrix format) that satisfy the known properties can be found.  
41 Selecting the right adjacency matrix among all of these suitable configurations or evaluating their relative  
42 probabilities may be done using different modeling tools (e.g. using the principle of maximum entropy or  
43 Bayesian inference). Second, as suggested by Strydom *et al.* (2021a), network structure can also be used to  
44 facilitate the estimation of pairwise species interactions. For example, by knowing a network's adjacency  
45 matrix (or the constrained space of feasible networks), we can decrease the volume of information we need to  
46 infer interactions (e.g. by knowing how many prey each species should have). Moreover, network structure can  
47 be used as validation to ensure that inferred interactions collectively satisfy known structural properties.  
48 Overall, partial knowledge of a network's properties can help us reconstruct its emerging structure and  
49 individual interactions when data is lacking.

50 Predicting ecological networks and understanding the biological mechanisms that shape species interactions are  
51 two complementary aims of network ecology. On one hand, prediction is the process of estimating the value of  
52 an unknown variable using available data and appropriate statistical models and assumptions. The goal of  
53 predictive models is to minimize predictive errors and provide reliable estimates of the unknown variable.  
54 These models have the potential to fill many data gaps about species interactions and network structure by  
55 making sound predictions and generating new data. A variety of such models have recently been developed  
56 using machine learning and other statistical tools, most of which are presented in Strydom *et al.* (2021a). On  
57 the other hand, statistical models can also be built to describe or understand a parameter or mechanism of  
58 interest. For example, null models are statistical models that generate baseline distributions to evaluate the  
59 significance of an observed measure. These distributions are obtained from different randomizations of a

network that maintain some of its properties while deliberately omitting others. By assessing the deviation of empirical data from these reference distributions, we can evaluate the degree to which the processes underlying the null model can accurately capture the observed measures, providing a benchmark to assess the importance of additional factors such as the omitted properties (Fortuna & Bascompte (2006), Delmas *et al.* (2019)). This makes null models a valuable tool to identify the ecological mechanisms that drive species interactions and constrain the structure of ecological networks. The distinction between understanding and predicting is crucial when using statistical and mathematical models as this can impact how we use and interpret them. Numerous predictive and null models have been used to analyze ecological networks (Delmas *et al.* (2019), Strydom *et al.* (2021a)). Nevertheless, given the constrained nature of ecological networks, it is surprising that the principle of maximum entropy, a mathematical method designed for both the description and prediction of constrained systems, has been barely used in network ecology.

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

72 The principle of maximum entropy (MaxEnt) is a mathematical method of finding probability distributions, strongly rooted in statistical mechanics and information theory (Jaynes (1957a), Jaynes (1957b), Harremoës & Topsøe (2001)). Starting from a set of constraints given by prior knowledge of a system (i.e. what we call state variables), this method helps us find least-biased probability distributions subject to the constraints. These probability distributions are guaranteed to be unique given our prior knowledge and represent the most we can say about a system without making more assumptions. For example, if we know the number of species and total number of individuals in a biological community (our state variables), we can calculate the average number of individuals per species (our constraint). Using the principle of maximum entropy, one could show that the least-biased species abundance distribution (i.e. the distribution of the number of individuals of each species), constrained by the average species abundance, follows an exponential distribution whose parameter is determined by the values of the state variables (Frank & Smith (2011), Harte & Newman (2014)). However, this does not imply that this distribution will be the best fit to empirical data. The challenge is to find the right set of constraints that would best reproduce distributions found in nature.

85 MaxEnt states that the least-biased probability distribution given the constraints used is the one with the highest entropy (a measure of the uncertainty of a random process) among all probability distributions that satisfy these constraints. Many measures of entropy have been developed in physics (Beck (2009)), but only a fraction of them could be used as an optimization measure with the principle of maximum entropy. According to Beck

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

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

93 satisfies all of these properties. Finding the probability distribution  $p(n)$  that maximizes  $H$  under a set of  $m$   
94 constraints  $g$  can be done using the method of Lagrange multipliers. These constraints can include one or many  
95 properties of the probability distribution (e.g. its mean, variance, and range). However, the normalization  
96 constraint always needs to be included in  $g$  to make sure that  $p(n)$  sums to 1. The objective is then to find the  
97 values of the Lagrange multipliers  $\lambda_i$  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 Shannon's  
99 entropy to which we added terms that each sums to zero ( $g_i = c_i$ ).  $F$  is maximized by setting to 0 its partial  
100 derivative with respect to  $p(n)$ .

101 The principle of maximum entropy has been used in a wide range of disciplines, from thermodynamics,  
102 chemistry and biology (Martyushev & Seleznev (2006)) to graph and network theory (e.g. Park & Newman  
103 (2004), van der Hoorn *et al.* (2018)). MaxEnt has also been proven useful in ecology, e.g. in species distribution  
104 (Phillips *et al.* (2006)) and macroecological (Harte *et al.* (2008), Harte & Newman (2014)) models. In network  
105 ecology, it can be used to generate null models of network structure with either fixed or fluctuating constraints  
106 (Caruso *et al.* (2022)) and predict important network properties using limited data. For example, it has been  
107 used to predict the degree distribution of bipartite networks from the number of species and the number of  
108 interactions (Williams (2011)) and the strength of biotic interactions from species relative abundances (Stock *et*  
109 *al.* (2021)). However, to the best of our knowledge, MaxEnt has never been used to predict food-web structure  
110 directly, even though food webs are among the most documented and widespread ecological networks (Ings *et*  
111 *al.* (2009)).

112 Food-web properties that can be derived using MaxEnt are varied and pertain to different elements of the

network (i.e. at the species, interaction and community levels). Because MaxEnt is a method of finding least-biased probability distributions given partial knowledge about a system, these properties need to be represented probabilistically. For example, at the species level, MaxEnt can be used to predict the distribution of trophic levels among species, as well as the distribution of species' vulnerability (number of predators) and generality (number of prey). By contrast, at the interaction level, predictions can be made on the distribution of interaction strengths in weighted food webs. At the community level, it can generate probability distributions of many measures of a network's emerging structure and of networks themselves (i.e. a probability distribution governing the occurrence of different network configurations). Because the decomposition of a network's adjacency matrix into a product of matrices can yield a vector of relative values (e.g. a vector of relative singular values), MaxEnt can also be used to find the configuration whose distribution of relative values is of maximum entropy. This configuration would be the one with the greatest entropy (or complexity, Strydom *et al.* (2021b)) among all configurations that satisfy the constraints. Overall, the potential of this method in the study of food webs is broad. The applicability and performance of MaxEnt mostly depend on the ecological information available and on our capacity to find the right set of state variables that best represent natural systems and to translate them into appropriate statistical constraints. Having a validated maximum entropy model for the system at hand allows us to make least-biased predictions using a minimal amount of data, as well as identify the most important ecological processes shaping that system. In other words, MaxEnt can help us better understand and predict the structure of ecological networks worldwide.

## 131 Analytical and heuristic approaches

132 In this contribution, we used two complementary approaches to predict the structure of food webs using the principle of maximum entropy. The first approach consists in deriving constrained probability distributions of 133 given network properties analytically, whereas the second approach consists in finding the adjacency matrix of 134 maximum entropy heuristically, from which network properties can be measured. We compared our predictions 135 against empirical data and null and neutral models commonly used in network ecology. We focus on 136 deterministic (non-probabilistic) and unweighted (Boolean) food webs in both approaches for two reasons. (1) 137 Most sampled food webs lack estimates of interaction probabilities (e.g. spatial variability of interactions) and 138 strengths (e.g. energy fluxes between species) because of inherent difficulties in measurement. (2) The methods 139 and measures to analyze the structure of binary food webs are more developed and tested (Delmas *et al.* 140 (2019)), which increases the range of analysis we can perform and simplifies their interpretation. However, our 141

142 framework can be applied to all types of ecological networks and a wider variety of measures.

143 For the first approach (analytic), we focus on species-level properties. Specifically, we derived the joint degree  
144 distribution (i.e. the joint probability distribution that a species has a given number of prey and predators in its  
145 network) of maximum entropy using only the number of species  $S$  and the number of interactions  $L$  as state  
146 variables. Then, we calculated the degree distribution of maximum entropy directly from the joint degree  
147 distribution by summing probabilities over the joint degree distribution (since the degree of a species is its total  
148 number of prey and predators combined). Because of the scarcity of empirical data on the number of  
149 interactions in food webs, we also present a method to predict  $L$  from  $S$  (Box 1), thus allowing the prediction of  
150 the joint degree distribution from  $S$  solely.

151 For the second approach (heuristic), we focus on community-level properties. We used a flexible and heuristic  
152 model based on simulated annealing (an optimization algorithm, Kirkpatrick *et al.* (1983)) to find the network  
153 configuration *close* to maximum entropy and measured its structure. We developed this heuristic model because  
154 the analytical derivation of a maximum entropy graph model of food webs is difficult, and because this model is  
155 readily applicable to other types of ecological networks and measures. Indeed, the mathematical representation  
156 of food webs (i.e. directed simple graphs frequently having self-loops) makes the optimization of maximum  
157 entropy graph models more complicated than with many other types of (non-ecological) networks. In other  
158 words, deriving a probability distribution on the graphs themselves is difficult when working with food webs.  
159 We built two types of heuristic MaxEnt models depending on the constraint used. Our type I MaxEnt model  
160 uses the connectance of the network (i.e. the ratio  $L/S^2$ ) as a constraint, whereas our type II MaxEnt model uses  
161 the whole joint degree sequence as a constraint.

## 162 Methods

### 163 Analytical maximum entropy models

164 The analytical approach is the most common way to use and develop maximum entropy models. As shown  
165 above, starting from a defined set of constraints, a mathematical expression for the target distribution is derived  
166 using the method of Lagrange multipliers. The derived distribution of maximum entropy is unique and least  
167 biased given the constraints used. Although we refer to this approach as analytic, finding the values of the  
168 Lagrange multipliers usually requires the use of numerical methods. Here we use MaxEnt to derive two

169 species-level properties in food webs: the joint degree distribution and the degree distribution. The degree  
170 distribution has driven the attention of ecologists because of its role in determining the assembly of ecological  
171 networks (Vázquez (2005)), shaping their emerging structure (Fortuna *et al.* (2010)), and understanding  
172 interaction mechanisms (Williams (2011)). As noted above, although the degree distribution of maximum  
173 entropy has already been derived in bipartite networks (Williams (2011)), we show in much greater detail its  
174 mathematical derivation in food webs. But first, we derived the joint degree distribution, a related property that  
175 holds significantly more ecological information than the degree distribution.

176 We tested our analytical MaxEnt models against open food-web data queried from three different sources and  
177 integrated into what we call our *complete dataset*. These sources include (1) terrestrial and aquatic food webs  
178 sampled globally and archived on the ecological interactions database Mangal; (2) freshwater stream food webs  
179 from New Zealand; and (3) aquatic food webs from Tuesday Lake (Michigan, United States). First, all food  
180 webs archived on `mangal.io` (Poisot *et al.* (2016), Banville *et al.* (2021)) were directly queried from the  
181 database ( $N = 235$ ). Most ecological networks archived on Mangal are multilayer networks, i.e. networks that  
182 describe different types of interactions. We kept all networks whose interactions were mainly of predation and  
183 herbivory types and removed the largest network ( $S = 714$ ) for computational efficiency reasons. Then, to this  
184 set, we added food webs from two different sources: the New Zealand dataset ( $N = 21$ , Pomeranz *et al.* (2018))  
185 and the Tuesday Lake dataset ( $N = 2$ , Cohen *et al.* (2003)). Our complete dataset thus contained a total of 257  
186 food webs. These complex food webs differ in their level of resolution and sampling effort, which may introduce  
187 noise in the estimation of their properties, especially given their large number of interacting elements. However,  
188 because our MaxEnt models are applied on imperfect data, they aim at reproducing the *sampled* structure of  
189 food webs, not their actual structure. All code and data to reproduce this article are available in the Open  
190 Science Framework (OSF.IO/KT4GS). Data cleaning, simulations and analyses were conducted in Julia v1.8.0.

## 191 **Joint degree distribution**

192 The joint degree distribution  $p(k_{in}, k_{out})$  of a food web with  $S$  species is a joint discrete probability distribution  
193 describing the probability that a species has  $k_{in}$  predators and  $k_{out}$  prey, with  $k_{in}$  and  $k_{out} \in [0, S]$ . Basal species  
194 (e.g. plants) have a  $k_{out}$  of 0, whereas top predators have a  $k_{in}$  of 0. In contrast, the maximum number of prey  
195 and predators a species can have is set by the total number of species in the food web. Here we show how the  
196 joint degree distribution of maximum entropy can be obtained given knowledge of the number of species  $S$  and  
197 the number of interactions  $L$ .

198 We want to maximize Shannon's entropy

$$H = - \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S p(k_{in}, k_{out}) \log p(k_{in}, k_{out}) \quad (3)$$

199 subject to the following constraints:

$$g_1 = \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S p(k_{in}, k_{out}) = 1; \quad (4)$$

$$g_2 = \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S k_{in} p(k_{in}, k_{out}) = \langle k_{in} \rangle = \frac{L}{S}; \quad (5)$$

$$g_3 = \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S k_{out} p(k_{in}, k_{out}) = \langle k_{out} \rangle = \frac{L}{S}. \quad (6)$$

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

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

$$F = H - \lambda_1(g_1 - 1) - \lambda_2 \left( g_2 - \frac{L}{S} \right) - \lambda_3 \left( g_3 - \frac{L}{S} \right), \quad (7)$$

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

$$\frac{\partial H}{\partial p(k_{in}, k_{out})} = \lambda_1 \frac{\partial g_1}{\partial p(k_{in}, k_{out})} + \lambda_2 \frac{\partial g_2}{\partial p(k_{in}, k_{out})} + \lambda_3 \frac{\partial g_3}{\partial p(k_{in}, k_{out})}. \quad (8)$$

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

$$-\log p(k_{in}, k_{out}) - 1 = \lambda_1 + \lambda_2 k_{in} + \lambda_3 k_{out}. \quad (9)$$

213 Then, solving eq. 9 for  $p(k_{in}, k_{out})$ , we obtain:

$$p(k_{in}, k_{out}) = \frac{e^{-\lambda_2 k_{in} - \lambda_3 k_{out}}}{Z}, \quad (10)$$

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

$$Z = \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S e^{-\lambda_2 k_{in} - \lambda_3 k_{out}}. \quad (11)$$

216 After substituting  $p(k_{in}, k_{out})$  in eq. 5 and eq. 6, we get a nonlinear system of two equations and two unknowns:

$$\frac{1}{Z} \sum_{k_{in}=0}^S \sum_{k_{out}=0}^S k_{in} e^{-\lambda_2 k_{in} - \lambda_3 k_{out}} = \frac{L}{S}; \quad (12)$$

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

217 We solved eq. 12 and eq. 13 numerically using the Julia library JuMP.jl v0.21.8 (Dunning *et al.* (2017)).

218 JuMP.jl supports nonlinear optimization problems by providing exact second derivatives that increase the  
219 accuracy and performance of its solvers. The estimated values of  $\lambda_2$  and  $\lambda_3$  can be substituted in eq. 10 to have  
220 a more workable expression for the joint degree distribution.

221 We assessed the empirical support of this expression using all food webs in our complete dataset. First, we  
222 predicted the joint degree distribution of maximum entropy for each of these food webs, i.e. using their number  
223 of species and number of interactions as state variables. Then, we sampled one realization of the joint degree  
224 sequence for each network using the probabilities given by the joint degree distribution, while fixing the total  
225 number of interactions. This gave us a random realization of the number of prey and predators for each node in  
226 each network. We standardized the predicted  $k_{out}$  and  $k_{in}$  by the total number of species in their network to  
227 generate relative values, which can be compared across networks.

228 **Degree distribution**

229 The joint degree distribution derived using MaxEnt (eq. 10) can be used to obtain the degree distribution of  
230 maximum entropy. The degree distribution  $p(k)$  represents the probability that a species has  $k$  interactions in its  
231 food web, with  $k = k_{in} + k_{out}$ . It can thus be obtained from the joint degree distribution as follows:

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

232 The degree distribution can also be obtained directly using the principle of maximum entropy, as discussed in  
233 Williams (2011). This gives the following distribution:

$$p(k) = \frac{e^{-\lambda_2 k}}{Z}, \quad (14)$$

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

235 This can be solved numerically using the constraint of the average degree  $\langle k \rangle = \frac{2L}{S}$  of a species, yielding an  
236 identical solution to the one obtained using the joint degree distribution as an intermediate.

$$\frac{1}{Z} \sum_{k=0}^S k e^{-\lambda_2 k} = \frac{2L}{S} \quad (15)$$

237 Note that the mean degree is twice the value of the linkage density because every interaction must be counted  
238 twice when we add in and out-degrees together.

239 **Heuristic maximum entropy models**

240 With the analytical approach, we showed how important measures of food-web structure (e.g. the degree  
241 distribution and the joint degree distribution) can be derived with the principle of maximum entropy using  
242 minimal knowledge about a biological community. This type of model, although useful to make least-biased  
243 predictions of many network properties, can be hard to apply for other measures. There are dozens of measures  
244 of network structure (Delmas *et al.* (2019)) and many are not directly calculated with mathematical equations  
245 but are instead estimated algorithmically. Moreover, the applicability of this method to empirical systems is  
246 limited by the state variables we can measure and use. Here, we suggest a more flexible method to predict many

247 measures of network structure simultaneously, i.e. by finding heuristically the network configuration of  
248 maximum SVD (singular value decomposition) entropy given partial knowledge of its emerging structure.

249 **From Shannon's to SVD entropy**

250 The principle of maximum entropy can be applied on the network itself if we decompose its adjacency matrix  
251 into a non-zero vector of relative values. This is a necessary step when working with food webs, which are  
252 frequently expressed as a matrix  $A = [a_{ij}]$  of Boolean values representing the presence ( $a_{ij} = 1$ ) or absence  
253 ( $a_{ij} = 0$ ) of an interaction between two species  $i$  and  $j$ . Knowing one or many properties of a food web of  
254 interest (e.g. its number of species and number of interactions), we can simulate its adjacency matrix randomly  
255 by using this known ecological information to constrain the space of potential networks. The entropy of this  
256 hypothetical matrix can then be measured after decomposing it into appropriate values. Simulating a series of  
257 networks until we find the one having the highest entropy allows us to search for the most complex food-web  
258 configuration given the ecological constraints used. This configuration is the least-biased one considering the  
259 information available. In other words, the most we can say about a network's adjacency matrix, without making  
260 more assumptions than the ones given by our incomplete knowledge of its structure, is the one of maximum  
261 entropy. Generating the most complex network that corresponds to this structure allows us to explore more  
262 easily other properties of food webs under MaxEnt.

263 Shannon's entropy can only be calculated on conventional probability distributions such as the joint degree  
264 distribution. This is an issue when working with the adjacency matrix of ecological networks. For this reason,  
265 we need to use another measure of entropy if we want to predict a network's configuration directly using  
266 MaxEnt. We used the SVD entropy as our measure of entropy, which is an application of Shannon's entropy to  
267 the relative non-zero singular values of a food web's adjacency matrix (Strydom *et al.* (2021b)). These values  
268 are obtained by performing singular value decomposition (SVD) on the matrix. SVD is a mathematical method  
269 that decomposes a matrix into three separate matrices, one of which contains singular values on the diagonal.  
270 We selected the non-zero singular values (i.e. the informative singular values that reflect a non-negligible  
271 contribution to the structure of the matrix) by keeping the largest  $R$  values, where  $R$  is the rank of the matrix  
272 (Golub *et al.* (1987)). We measured SVD entropy as follows:

$$J = - \sum_{i=1}^R s_i \log s_i, \quad (16)$$

273 where  $s_i$  are the relative singular values of the adjacency matrix ( $s_i = \sigma_i / \sum_{i=1}^R \sigma_i$ , where  $\sigma_i$  are the singular  
274 values). Note that the distribution of relative singular values is analogous to a probability distribution, with  
275  $0 < s_i < 1$  and  $\sum s_i = 1$ . This measure also satisfies all four properties of an appropriate entropy measure  
276 above-mentioned, while being a proper measure of the internal complexity of food webs (Strydom *et al.*  
277 (2021b)). Following Strydom *et al.* (2021b), we standardized it with the rank of the matrix (i.e.  $J / \ln(R)$ ) to  
278 account for the difference in dimensions between networks (*sensu* Pielou's evenness, Pielou (1975)).

## 279 **Types I and II heuristic MaxEnt models**

280 We used SVD entropy to predict the network configuration of maximum entropy (i.e. of maximum complexity)  
281 heuristically given different constraints for all food webs in our complete dataset. We built two types of heuristic  
282 MaxEnt models that differ on the constraint used. The type I heuristic MaxEnt model is based on connectance,  
283 whereas the type II heuristic MaxEnt model is based on the joint degree sequence. These models are thus based  
284 on the same constraints as the types I (Fortuna & Bascompte (2006)) and II (Bascompte *et al.* (2003)) null  
285 models (Box 2) frequently used to generate random networks topologically. This allows a direct comparison of  
286 the performance of null and heuristic MaxEnt models in reproducing the emerging structure of empirical food  
287 webs.

288 For each network, we estimated their configuration of maximum entropy given each of these constraints. For  
289 both types of heuristic MaxEnt models, we used a simulated annealing algorithm with 4 chains, 2000 steps and  
290 an initial temperature of 0.2. For each food web, we first generated one random Boolean matrix per chain while  
291 fixing the number of species. We also maintained the total number of interactions (i.e. the sum of all elements  
292 in the matrix) in the type I MaxEnt model and the joint degree sequence (i.e. the rows and columns sums) in the  
293 type II MaxEnt model. These were our initial configurations. Then, we swapped interactions sequentially while  
294 maintaining the original connectance or joint degree sequence. Configurations with a higher SVD entropy than  
295 the previous one in the chain were always accepted, whereas they were accepted with a probability conditional  
296 to a decreasing temperature and the difference in SVD entropy when lower. The final configuration with the  
297 highest SVD entropy among the four chains constitutes the estimated maximum entropy configuration of a food  
298 web given the constraint used.

299 **Structure of MaxEnt food webs**

300 We measured various properties of these configurations of maximum entropy to evaluate how well they  
301 reproduce the structure of sampled food webs. Specifically, we evaluated their nestedness  $\rho$ , their maximum  
302 trophic level  $maxtl$ , their network diameter  $diam$ , their average maximum similarity between species pairs  
303  $MxSim$  (Williams & Martinez (2000)), their proportion of cannibal species *Cannib*, their proportion of  
304 omnivorous species *Omniv*, their SVD entropy, and their motif profile. Nestedness indicates how much the diet  
305 of specialist species is a subset of the one of generalists (Delmas *et al.* (2019)) and was measured using the  
306 spectral radius of the adjacency matrix (Staniczenko *et al.* (2013)). Next, the network diameter represents the  
307 longest of the shortest paths between all species pair (Albert & Barabasi (2002)). Further, cannibal species are  
308 species that can eat individuals of their own species (i.e. species having self-loops), whereas omnivorous  
309 species can prey on different trophic levels (Williams & Martinez (2000)). Finally, motifs are unique n-species  
310 connected subgraphs that can be considered simple building blocks of ecological networks (Milo *et al.* (2002),  
311 Stouffer *et al.* (2007)). There are 13 possible three-species motifs in food webs, including 5 with only single  
312 links ( $i \rightarrow j$ ) and 8 with double links ( $i \leftrightarrow j$ ), labeled S1-S5 and D1-D8 by Stouffer *et al.* (2007), respectively. A  
313 motif profile represents the proportion of each of these motifs in a network (i.e. their relative frequencies). All  
314 of these properties are relatively easy to measure and were chosen based on their ecological importance and  
315 prevalent use in network ecology (McCann (2011), Delmas *et al.* (2019)).

316 We compared the performance of both heuristic MaxEnt models in predicting these measures to the one of the  
317 null and neutral models (Box 2). We conducted these comparisons using two different datasets: (1) our  
318 complete dataset including most food webs archived on Mangal, as well as all food webs in the New Zealand  
319 and Tuesday Lake datasets, and (2) our *abundance dataset*, a subset of the complete dataset comprising all food  
320 webs having data on their species' relative abundances ( $N = 19$ ). Indeed, of the New Zealand and Tuesday Lake  
321 datasets, 19 networks had data on species' relative abundances that were used in the neutral model to better  
322 assess the performance of our heuristic models. We compared our models' predictions using these two datasets  
323 separately to minimize biases and to better represent food webs with abundance data.

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

325 Our analytical MaxEnt models require information on the number of species and the number of interactions.  
326 However, since the latter is rarely measured empirically, ecologists might need to use another predictive model

327 to estimate the total number of interactions in a food web before using MaxEnt. Here we show how this can be  
328 done by combining both models sequentially.

329 We used the flexible links model of MacDonald *et al.* (2020) to predict the number of interactions from the  
330 number of species in food webs. The flexible links model has been shown to make reliable predictions of  $L$   
331 while taking into account meaningful ecological constraints on its lower and upper bounds. Specifically,  
332 recognizing that there is a minimum of  $S - 1$  (if no isolated species) and a maximum of  $S^2$  (if all species  
333 interact) interactions in food webs, it estimates the number of the  $S^2 - (S - 1)$  *flexible links* that are realized.  
334 This represents the number of realized interactions above the minimum (i.e.  $L_{FL} = L - (S - 1)$ ). The  
335 probability that a flexible link is realized is treated as constant within a particular food web but differs across  
336 food webs to capture variations in connectance. Therefore, assuming that flexible links are independent  
337 Bernoulli events, the total number of flexible links that are realized in a food web can be obtained from a  
338 beta-binomial distribution BB with  $S^2 - (S - 1)$  trials and parameters  $\alpha = \mu e^\phi$  and  $\beta = (1 - \mu)e^\phi$ :

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

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

341 We fitted the flexible links model on all food webs in our complete dataset and estimated the parameters of  
342 eq. 17 using a Hamiltonian Monte Carlo sampler with static trajectory (4 chains and 3000 iterations):

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

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

347 The flexible links model is a generative model, i.e. it can generate plausible values of the predicted variable. We  
348 thus simulated 1000 values of  $L$  for different values of  $S$  using the joint posterior distribution of our model  
349 parameters (eq. 18) and calculated the mean degree for each simulated value. The resulting distributions are  
350 shown in the left panel of fig. 1 for three different values of species richness. In the right panel of fig. 1, we

351 show how the probability distribution for the mean degree constraints can be used to generate a distribution of  
352 maximum entropy degree distributions since each simulated value of mean degree generates a different  
353 maximum entropy degree distribution (eq. 14 and eq. 15).

354 [Figure 1 about here.]

355 **Fig 1. Maximum entropy degree distributions with predicted numbers of interactions.** (a) Probability  
356 density of the mean degree of a food web obtained using different values of species richness  $S$ . The number of  
357 interactions  $L$  was simulated 1000 times using the flexible links model fitted to all empirical networks (eq. 18).  
358 The mean degrees  $2L/S$  were then obtained from these simulated values. (b) Degree distributions of maximum  
359 entropy for a network of  $S = 27$  species and different numbers of interactions. The numbers of interactions  
360 correspond to the lower and upper bounds of the 67, 89, and 97 percentile intervals (PI), as well as the median  
361 (in blue), of the counterfactuals of the flexible links model. Each degree distribution of maximum entropy was  
362 obtained using eq. 14 after solving numerically eq. 15 using different values of the mean degree constraint.

## 363 **Box 2 - Corresponding null and neutral models**

### 364 **Null models (types I and II)**

365 The predictions of our heuristic maximum entropy models were compared against two topological null models.  
366 These null models use the same ecological information as our heuristic models and thus constitute an adequate  
367 baseline for comparison. The first is the type I null model of Fortuna & Bascompte (2006), in which the  
368 probability that a species  $i$  predares on another species  $j$  is given by

$$369 p(i \rightarrow j) = \frac{L}{S^2}. \quad (19)$$

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

$$372 p(i \rightarrow j) = \frac{1}{2} \left( \frac{k_{in}(j)}{S} + \frac{k_{out}(i)}{S} \right), \quad (20)$$

371 where  $k_{in}(j)$  and  $k_{out}(i)$  are the in and out-degrees of species  $j$  and  $i$ , respectively. The type I null model is  
372 based on connectance, whereas the type II null model is based on the joint degree sequence. Therefore, the type  
373 I and II topological null models correspond to our type I and II heuristic MaxEnt models, respectively, since  
374 they use similar constraints.

375 We generated probabilistic networks using both types of null models for all empirical food webs in our complete  
376 dataset. Then, we converted these networks to adjacency matrices of Boolean values by generating 100 random  
377 networks for each of these probabilistic webs and kept the  $L$  entries that were sampled the most amount of  
378 times, with  $L$  given by the number of interactions in each food web. This ensured that the resulting null  
379 networks had the same number of interactions as their empirical counterparts. Thus, for each null model, we  
380 ended up with one null adjacency matrix for each empirical network.

## 381 Neutral model

382 We also compared our heuristic MaxEnt models to a neutral model of relative abundances, in which the  
383 probabilities of interaction are given by

$$384 p(i \rightarrow j) \propto \frac{n_i}{N} \times \frac{n_j}{N}, \quad (21)$$

385 where  $n_i$  and  $n_j$  are the abundances (or biomass) of both species and  $N$  is the total abundance (or biomass) of all  
386 species in the network. We generated neutral abundance matrices for all empirical food webs in our abundance  
387 dataset and converted these weighted networks to adjacency matrices of Boolean values using the same method  
388 as the one we used for our null models.

## 388 Results and Discussion

### 389 Analytical maximum entropy models

390 We first discuss the predictive capacity of our analytical models. The relationship between the relative numbers  
391 of prey  $k_{out}$  and predators  $k_{in}$  in empirical networks and obtained from the joint degree distributions of  
392 maximum entropy is depicted in the left and central panels of fig. 2, respectively. We observe that our analytical  
393 model predicts higher values of generality and vulnerability compared to empirical food webs (i.e. relative

394 values of  $k_{out}$  and  $k_{in}$  both closer to 1) for many species. In other words, our model predicts that species that  
395 have many predators also have more prey than what is observed empirically (and conversely). This is not  
396 surprising, given that our model did not include biological factors preventing generalist predators from having  
397 many prey. Nevertheless, with the exception of these generalist species, MaxEnt adequately predicts that most  
398 species have low generality and vulnerability values.

399 [Figure 2 about here.]

400 **Fig 2. Prediction errors of the relative number of predators and prey.** The relative number of predators  
401 ( $k_{in}$ ) is plotted against the relative number of prey ( $k_{out}$ ) for each species in all (a) empirical and (b) predicted  
402 joint degree sequences. The predicted joint degree sequences were obtained after sampling one realization of  
403 the joint degree distribution of maximum entropy for each network while keeping the total number of  
404 interactions constant. (c) Difference between predicted and empirical values when species are ordered  
405 according to their total degree. Due to significant data overlap, all relationships are represented as 2D  
406 histograms. The color bar indicates the number of species that fall within each bin.

407 Examining the difference between predicted and empirical values for each species gives a slightly different  
408 perspective (right panel of fig. 2). To make that comparison, we must first associate each of our predictions with  
409 a specific species in a network. Indeed, our predicted joint degree sequences have the same number of species  
410 (elements) as their empirical counterparts, but they are species agnostic. In other words, instead of predicting a  
411 pair of values for each species directly (i.e. the number of prey and predators of a given species  $i$ ), we predicted  
412 the entire joint degree sequence without taking into account species' identity (i.e. the distribution of the number  
413 of prey and predators for the entire set of species, without knowing which values belong to which species). The  
414 challenge is thus to adequately associate predictions with empirical data. In fig. 2, we present these differences  
415 when species are ordered by their total degree in their respective networks (i.e. by the sum of their in and  
416 out-degrees). This means that the species with the highest total degree in its network will be associated with the  
417 highest prediction, and so forth. Doing so, we see that species predicted to have a higher number of predators  
418 than what is observed generally have a lower number of prey than what is observed (and conversely). This is  
419 also shown in S1 Fig, which represents the relationship between prediction errors in the *absolute* (non-relative)  
420 values of  $k_{out}$  and  $k_{in}$  across networks of varying levels of species richness. This is because the difference in  
421 total degree ( $k_{out} + k_{in}$ ) between predictions and empirical data is minimized when species are ranked by their  
422 total degree (i.e. the average deviation of the sum of relative  $k_{out}$  and  $k_{in}$  is close to 0 across all species). This

423 result thus shows that the difference between predicted and empirical total degrees is low for most species when  
424 ordered by their total degrees. There are no apparent biases towards in or out degrees. In S2 Fig, we show how  
425 these differences change when species are instead ordered by their out-degrees (left panel) and in-degrees (right  
426 panel), i.e. when minimizing the error in the estimation of the out and in-degrees, respectively.

427 Another way to evaluate the empirical support of the sampled joint degree sequences is to compare their shape  
428 with the ones of empirical food webs. We described the shape of a joint degree sequence by measuring the  
429 distance between its in and out-degree sequences (i.e. the distance between its marginal distributions). To do so,  
430 we calculated the Kullback–Leibler (KL) divergence (Kullback & Leibler (1951)) between the in and out-degree  
431 sequences of each predicted and empirical distribution. The KL divergence is a measure of relative entropy  
432 describing the difference between two distributions. Low values indicate high similarity between the in and  
433 out-degree sequences and suggest that the joint degree sequence has a high level of symmetry. We compared the  
434 shape of the empirical and predicted joint degree sequences in the left panel of fig. 3. As expected, our model  
435 predicts more similar in-degree and out-degree sequences than empirical data (shown by lower KL divergence  
436 values). However, the difference between the KL divergence of predicted and empirical joint degree sequences  
437 decreases with connectance (right panel of fig. 3). This might be because food webs with a low connectance are  
438 harder to predict than food webs with a high connectance. Indeed, in low connectance systems, what makes two  
439 species interact may be more important for prediction than in high connectance systems, in which what prevents  
440 species from interacting may be more meaningful. This implies that more ecological information may be needed  
441 in food webs with a low connectance because more ecological processes determine interactions compared to  
442 non-interactions. Therefore, other ecological constraints might be needed to account for the asymmetry of the  
443 joint degree distribution, especially for networks with a lower connectance. Nevertheless, our MaxEnt model  
444 seems to capture quite well the shape of the joint degree sequence for networks having a high connectance.

445 [Figure 3 about here.]

446 **Fig 3. Shape of empirical and predicted joint degree sequences.** (a) Probability density of KL divergence  
447 between in and out-degree sequences of empirical and predicted joint degree sequences. (b) Difference between  
448 the KL divergence of empirical and predicted joint degree sequences as a function of connectance. The  
449 predicted joint degree sequences were obtained after sampling one realization of the joint degree distribution of  
450 maximum entropy for each network while keeping the total number of interactions constant.

451 Regarding the degree distribution of maximum entropy, one aspect that informs us of its ecological realism is

452 the number of isolated species it predicts. As MacDonald *et al.* (2020) pointed out, the size of food webs should  
453 at least be of  $S - 1$  interactions, since a lower number would yield isolated species, i.e. species without any  
454 predator or prey. Because non-basal species must eat to survive, isolated species could indicate that other  
455 species are missing; otherwise, isolated species should be removed from the network. In S3 Fig, we show that  
456 the degree distribution of maximum entropy, given  $S$  and  $L$ , gives a very low probability that a species will be  
457 isolated in its food web (i.e. having  $k = 0$ ) when  $L > S - 1$ . However, under our purely information-theoretic  
458 model, the probability that a species is isolated is quite high when the total number of interactions is below  $S - 1$ .  
459 Moreover, the expected proportion of isolated species rapidly declines by orders of magnitude with increasing  
460 numbers of species and interactions. This supports the ecological realism of the degree distribution of  
461 maximum entropy derived above. Nevertheless, ecologists wanting to model a system without allowing isolated  
462 species could simply change the lower limit of  $k$  to 1 in eq. 15 and solve the resulting equation numerically.

#### 463 Heuristic maximum entropy models

464 In this section, we explore the predictions of our heuristic models. Overall, we found that the models based on  
465 the joint degree sequence (i.e. the type II null and heuristic MaxEnt models) reproduced the structure of  
466 empirical food webs much better than the ones based on connectance (i.e. the type I null and heuristic MaxEnt  
467 models, tbl. 1). This suggests that the predictive capacity of connectance might be more limited than what was  
468 previously suggested (Poisot & Gravel (2014)). On the other hand, the neutral model of relative abundances was  
469 surprisingly good at predicting the maximum trophic level and the network diameter (tbl. 2). However, with the  
470 exception of the network diameter, the type II heuristic MaxEnt model was better at predicting network  
471 structure than the neutral model for most measures considered. This might be because, although neutral  
472 processes are important, they act in concert with niche processes in determining species interactions  
473 (Bartomeus *et al.* (2016), Canard *et al.* (2014), Poisot *et al.* (2015), Pomeranz *et al.* (2019)). The joint degree  
474 sequence captures information on both neutral and niche processes because the number of prey and predators a  
475 species has is determined by its relative abundance and biological traits. These results thus show that having  
476 information on the number of prey and predators for each species substantially improves the prediction of  
477 food-web structure, both compared to models solely based on connectance and to the ones solely based on  
478 species relative abundances.

**Table 1: Standardized mean difference between predicted network measures and empirical data for all food webs in our complete dataset ( $N = 257$ ).** Positive (negative) values indicate that the measure is overestimated (underestimated) on average. Null 1: Type I null model based on connectance. MaxEnt 1: Type I heuristic MaxEnt model based on connectance. Null 2: Type II null model based on the joint degree sequence. MaxEnt 2: Type II heuristic MaxEnt 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 difference between predicted network measures and empirical data for all food webs in our abundance dataset ( $N = 19$ ).** Positive (negative) values indicate that the measure is overestimated (underestimated) on average. Neutral: Neutral model of relative abundances. Null 1: Type I null model based on connectance. MaxEnt 1: Type I heuristic MaxEnt model based on connectance. Null 2: Type II null model based on the joint degree sequence. MaxEnt 2: Type II heuristic MaxEnt 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

479 Next, the predictions of the type II heuristic MaxEnt model can be compared to its null model counterpart. On  
480 average, the type II heuristic MaxEnt model was better at predicting nestedness ( $0.62 \pm 0.08$ ) than its  
481 corresponding null model ( $0.73 \pm 0.05$ ; empirical networks:  $0.63 \pm 0.09$ ) for networks in our complete dataset  
482 (tbl. 1). This might in part be due to the fact that nestedness was calculated using the spectral radius of the  
483 adjacency matrix, which directly leverages information on the network itself just like the heuristic MaxEnt

model. The proportion of self-loops (cannibal species) was also better predicted by the type II heuristic MaxEnt model in comparison to the type II null model. However, the type II null model was better at predicting network diameter and average maximum similarity between species pairs, and predictions of the maximum trophic level and the proportion of omnivorous species were similar between both types of models. We believe that this is because increasing the complexity of a food web might increase its average and maximum food-chain lengths. In comparison, the null model was more stochastic and does not necessarily produce more complex food webs with longer food-chain lengths.

Moreover, we found that the entropy of empirical food webs was slightly lower than their maximum entropy when constrained by their joint degree sequence (S4 Fig). Empirical food webs had an SVD entropy of  $0.89 \pm 0.04$ , compared to an SVD entropy of  $0.94 \pm 0.03$  for networks generated using the type II heuristic MaxEnt model. The relationship between the SVD entropy of empirical food webs and their maximum entropy is plotted in the last panel of fig. 4. The slight increase in entropy confirms that our method generated more complex networks. Even though we found that many measures of empirical networks are close to the ones of their maximum entropy configuration, the relatively low predictability of entropy itself may be indicative of additional constraints shaping food-web structure, especially for networks with low SVD entropy. Incorporating more constraints into the model could increase its capacity to generate networks with an adequate level of complexity, as shown by the decrease in predictive errors of entropy of the type II heuristic MaxEnt model compared to the one based on connectance (tbl. 1). Additionally, we found no clear relationship between the increase in SVD entropy and the number of species, the number of interactions, and connectance (S5 Fig). This suggests that our model captured the complexity of small and large networks on a similar level and that its capacity to reproduce food-web structure was unrelated to the order and size of the network. In other words, the gap in entropy between empirical food webs and their maximum entropy configuration may be the result of additional constraints that were not taken into account in the model, regardless of the number of species and the number of interactions.

[Figure 4 about here.]

**Fig 4. Relationship between the structure of empirical and maximum entropy food webs.** Maximum entropy networks were obtained using the type II heuristic MaxEnt model based on the joint degree sequence. (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

513 webs. The identity line is plotted in each panel.

514 A direct comparison of the structure of maximum entropy food webs constrained by the joint degree sequence  
515 with empirical data also supports the results depicted in [tbl. 1](#). In [fig. 4](#), we show how well empirical measures  
516 are predicted by the type II heuristic MaxEnt model. Following our previous results, we found that nestedness  
517 was very well predicted by our model. However, the model overestimated the maximum trophic level and  
518 network diameter, especially when the sampled food web had intermediate values of these measures. In [S6 Fig](#),  
519 we show that the pairwise relationships between the four measures in [fig. 4](#) and species richness in empirical  
520 food webs are similar (in magnitude and sign) to the ones found in food webs generated using the type II  
521 heuristic MaxEnt model. This indicates that the number of species in the network does not seem to impact the  
522 ability of the model to reproduce food-web structure.

523 Notwithstanding its difficulties in reproducing adequate measures of food-chain lengths, the type II heuristic  
524 MaxEnt model can predict surprisingly well the proportions of three-species motifs in empirical food webs.  
525 Motifs have been shown to be the backbone of complex ecological networks on which network structure is built  
526 and play a crucial role in community dynamics and assembly (Stouffer & Bascompte (2011)). Differences in  
527 motif profiles between an observed food web and null model-generated ones can unveil important ecological  
528 mechanisms that contribute to network structure (Stouffer *et al.* (2007)). In [fig. 5](#), we show that the motif profile  
529 of networks generated using the type II heuristic MaxEnt model accurately reproduced the one of empirical  
530 data. This model made significantly better predictions than the ones based on connectance and the type II null  
531 model based on the joint degree sequence. This is also shown in [fig. 6](#), which reveals that the relationships  
532 between the proportions of single-link motifs in empirical food webs are similar to the ones in networks  
533 generated using the type II heuristic MaxEnt model. This is in contrast with the type I null and MaxEnt models  
534 based on connectance, which produced opposite relationships than what was observed empirically. Our findings  
535 show that generating the most complex food web constrained by the joint degree sequence using maximum  
536 entropy does not alter the proportions of three-species motifs on the whole. This suggests that motif profiles  
537 may simply be a statistical attribute of food webs driven by the joint degree sequence. However, given the  
538 incapacity of our MaxEnt models to accurately predict food-chain lengths, the way motifs interconnect with  
539 each other may hold greater biological significance than the proportion of motifs itself.

540

[Figure 5 about here.]

541 **Fig 5. Proportions of single-link three-species motifs in empirical and predicted food webs. S1:** Tri-trophic

542 chain (a top predator feeds on a meso-predator which feeds on a basal prey). S2: Omnivory (a top predator  
543 feeds on a meso-predator and a basal prey). S3: Tri-trophic feeding loop (a cyclic three-species predator-prey  
544 system). S4: Apparent competition (a predator feeds on two prey). S5: Exploitative competition (two predators  
545 feed on the same prey). Null 1: Type I null model based on connectance. MaxEnt 1: Type I heuristic MaxEnt  
546 model based on connectance. Null 2: Type II null model based on the joint degree sequence. MaxEnt 2: Type II  
547 heuristic MaxEnt model based on the joint degree sequence. Boxplots display the median proportion of each  
548 motif in food webs (middle horizontal lines), as well as the first (bottom horizontal lines) and third (top  
549 horizontal lines) quartiles. Vertical lines encompass all data points that fall within 1.5 times the interquartile  
550 range from both quartiles, and dots are data points that fall outside this range. Only the single-link motifs S1-S5  
551 are shown given the scarcity of double-link motifs in most empirical and predicted networks.

552 [Figure 6 about here.]

553 **Fig 6. Pairwise relationships between the proportions of single-link three-species motifs in empirical and**  
554 **predicted food webs.** S1: Tri-trophic chain. S2: Omnivory. S4: Apparent competition. S5: Exploitative  
555 competition. Null 1: Type I null model based on connectance. MaxEnt 1: Type I heuristic MaxEnt model based  
556 on connectance. Null 2: Type II null model based on the joint degree sequence. MaxEnt 2: Type II heuristic  
557 MaxEnt model based on the joint degree sequence. Regression lines are plotted in each panel. Motif S3 is not  
558 shown because of its low proportion in most empirical and predicted networks.

559 One of the challenges in implementing and validating a maximum entropy model is to discover where its  
560 predictions break down. The results depicted in [tbl. 1](#) and [fig. 4](#) show that our type II heuristic MaxEnt model  
561 can capture many high-level properties of food webs, but does a poor job of capturing others. This suggests that,  
562 although the joint degree sequence is an important driver of food-web structure, other ecological constraints  
563 might be needed to account for some emerging food-web properties, especially the ones regarding food-chain  
564 lengths. Nevertheless, [fig. 5](#) and [fig. 6](#) show that this model can reproduce surprisingly well motif profiles, one  
565 of the most ecologically informative properties of food webs. This suggests that the emerging structure of food  
566 webs is mainly driven by their joint degree sequence, although higher-level properties might need to be included  
567 in the model to ensure that food-chain lengths fall within realistic values.

568 **Conclusion**

569 The principle of maximum entropy is a mathematical method of finding least-biased probability distributions  
570 that have some specified properties given by prior knowledge about a system. We first applied this conventional  
571 MaxEnt approach on food webs to predict species-level properties, namely the joint degree distribution and the  
572 degree distribution of maximum entropy given known numbers of species and interactions. We found that the  
573 joint degree distributions of maximum entropy had a similar shape to the ones of empirical food webs in  
574 high-connectance systems. However, these MaxEnt distributions were more symmetric than the ones of  
575 empirical food webs when connectance was low, which suggests that other constraints might be needed to  
576 improve these predictions in low-connectance systems. Then, we used a slightly different approach that aimed at  
577 finding heuristically the network configuration with the highest SVD entropy, i.e. whose vector of relative  
578 singular values has maximum entropy. This network of maximum entropy is the most complex, or random,  
579 given the specified structure. We found that the heuristic maximum entropy model based on connectance did  
580 not predict the structure of sampled food webs very well. However, the heuristic maximum entropy model based  
581 on the entire joint degree sequence, i.e. on the number of prey and predators for each species, gave more  
582 convincing results. Indeed, this model reproduced food-web structure surprisingly well, including the highly  
583 informative motif profile. Nevertheless, it was not able to predict realistic food-chain lengths.

584 Our results bring to the forefront the role of the joint degree distribution in shaping food-web structure. This  
585 echoes the work of Fortuna *et al.* (2010), who found that the degree distribution of ecological networks drives  
586 their emerging structure such as their nestedness and modularity. Several measures of food webs have been  
587 analyzed when studying the ecological consequences of network structure (McCann (2011), Delmas *et al.*  
588 (2019)). In fact, following Williams (2011), we believe that there is a lot more ecological information in the  
589 deviation between these properties in empirical systems and their maximum entropy configuration given a fixed  
590 joint degree sequence.

591 **Alternative MaxEnt models**

592 In this contribution, we used a method based on simulated annealing to find the network configuration with the  
593 highest SVD entropy while fixing some aspects of its structure. However, there are different ways to generate  
594 adjacency matrices using MaxEnt. Another technique, also based on simulated annealing, could begin by  
595 generating a food web randomly with fixed numbers of species and interactions and calculating its joint degree

596 distribution. Pairs of interactions could then be swapped sequentially until we minimize the divergence between  
597 the calculated joint degree distribution and the one of maximum entropy obtained analytically. In that case, this  
598 is the entropy of the joint degree distribution that would be maximized, not the one of the network's topology.  
599 To a certain extent, this method would bridge the gap between the analytical and heuristic approaches presented  
600 in this article. More research is needed to compare the quality of different methods in generating adjacency  
601 matrices of food webs using MaxEnt.

602 Maximum entropy graph models are another type of method that predicts a distribution of adjacency matrices  
603 under soft or hard constraints (e.g. Park & Newman (2004), Cimini *et al.* (2019)). Under hard constraints, every  
604 network with a non-zero probability exactly satisfies the constraints on its structure. This is in contrast with soft  
605 constraints, which require that networks satisfy them on average (i.e. many networks with a non-zero probability  
606 do not have the exact structure set by the constraints). Maximum entropy graph models are helpful because they  
607 can provide probability distributions for many network properties by measuring the structure of all adjacency  
608 matrices with a non-zero probability. However, we consider that our approach based on simulated annealing is  
609 more flexible and more computationally efficient. Indeed, many measures of food-web structure are hard to  
610 translate into mathematical constraints. Moreover, because food webs are directed networks that can have  
611 self-loops, it makes the mathematical derivation of maximum entropy graph models difficult. We believe that  
612 identifying heuristically what constrains the topology of food webs is a useful first step before attempting to  
613 derive the mathematical formulation of a maximum entropy graph model for food webs.

## 614 Applications

615 Our analytical and heuristic models can be applied for different purposes. First, they could be used to generate  
616 first-order approximations of a network's properties when state variables are known empirically. For example,  
617 knowing the number of species in an ecological community, we can predict its number of interactions using the  
618 flexible links model and then predict its joint degree distribution with minimal biases using the principle of  
619 maximum entropy. This could prove particularly useful when predicting network structure at large spatial  
620 scales, subdividing the study area into smaller communities (e.g. grid cells). Indeed, because species richness  
621 and other ecological data are increasingly abundant (e.g. Dickinson *et al.* (2010)), validated MaxEnt models  
622 can be used to respond to a wider range of macroecological questions regarding food webs.

623 Second, our analytical model can be used to generate informative priors in Bayesian analyses of the structure of

624 ecological networks (e.g. Cirtwill *et al.* (2019)). Indeed, the probability distribution of maximum entropy  
625 derived using MaxEnt can be used as a prior that can be updated with novel data. For instance, if we know the  
626 number of species and the number of interactions, we can derive the degree distribution of maximum entropy,  
627 as shown in this contribution. The degree distribution represents the probability that a species can interact (as a  
628 predator or a prey) with a given number of other species. Data on species interactions can be used to update the  
629 prior degree distribution to generate a more accurate posterior distribution, thus improving our description and  
630 understanding of the system.

631 Third, our analytical and heuristic models can be used to make better predictions of pairwise species  
632 interactions by constraining the space of feasible networks, as discussed in Strydom *et al.* (2021a). In other  
633 words, we can use the network configuration or specific measures of food-web structure derived using MaxEnt  
634 to ensure that our predictions of interspecific interactions form feasible networks. This means that the  
635 probability that two species interact may be conditional on the network structure and the probability of  
636 interactions of all other species pairs. When data are limited, MaxEnt can be used to predict network structure  
637 on which pairwise probabilities of interactions are conditional.

638 Finally, our analytical and heuristic models can be used as alternative null models of ecological networks to  
639 better understand and identify the ecological processes driving food-web structure. Indeed, these mechanisms  
640 can be better described when analyzing the deviation of empirical data from MaxEnt predictions (Caruso *et al.*  
641 (2022)). A strong deviation would indicate that ecological mechanisms not captured by the statistical  
642 constraints are at play for the system at hand. For instance, the incapacity of a MaxEnt model to reliably predict  
643 entropy may be a compelling indication of additional constraints shaping food-web structure. If deviations are  
644 systematic, the maximum entropy model might need to be revised to include appropriate ecological constraints.  
645 This revision process helps us reflect on and identify what constrains food-web structure. However, it is  
646 important to note that tangible ecological mechanisms cannot be directly inferred from statistical distributions  
647 (Warren II *et al.* (2022)). Instead, by identifying the constraints of a system and by analyzing empirical  
648 deviations from maximum entropy predictions, MaxEnt can only help us redirect research efforts toward  
649 understanding the biological mechanisms behind these constraints.

650 The principle of maximum entropy can thus be applied to both the prediction and understanding of natural  
651 systems. The model's interpretation depends on how we use it. It can be used as a baseline distribution to  
652 identify the ecological processes organizing natural systems. It can also be used to generate predictions of  
653 ecological networks. This distinction between understanding and predicting is important when using and

654 interpreting MaxEnt models.

655 **Final remarks**

656 One of the biggest challenges in using the principle of maximum entropy is to identify the set of state variables  
657 that best reproduce empirical data. We found that the number of species and the number of interactions are  
658 important state variables for the prediction of the joint degree distribution. Similarly, we found that the numbers  
659 of prey and predators for each species in a food web are important state variables for the prediction of the  
660 network configuration. However, our predictions overestimated the symmetry of the joint degree distribution for  
661 our analytical model and the maximum trophic level and network diameter for our heuristic model. We should  
662 thus continue to play the ecological detective to find these other topological constraints that would improve the  
663 predictions of MaxEnt models and help us understand better what drives food-web structure.

664 **Acknowledgments**

665 We acknowledge that this study was conducted on land within the traditional unceded territory of the Saint  
666 Lawrence Iroquoian, Anishinabewaki, Mohawk, Huron-Wendat, and Omàmiwininiwak nations.

667 **References**

- 668 Albert, R. & Barabasi, A.L. (2002). [Statistical mechanics of complex networks](#). *Reviews of Modern Physics*,  
669 74, 47–97.
- 670 Banville, F., Vissault, S. & Poisot, T. (2021). [Mangal.jl and EcologicalNetworks.jl: Two complementary](#)  
671 [packages for analyzing ecological networks in Julia](#). *Journal of Open Source Software*, 6, 2721.
- 672 Bartomeus, I., Gravel, D., Tylianakis, J.M., Aizen, M.A., Dickie, I.A. & Bernard-Verdier, M. (2016). [A](#)  
673 [common framework for identifying linkage rules across different types of interactions](#). *Functional Ecology*,  
674 30, 1894–1903.
- 675 Bascompte, J., Jordano, P., Melian, C.J. & Olesen, J.M. (2003). [The nested assembly of plant-animal](#)  
676 [mutualistic networks](#). *Proceedings of the National Academy of Sciences of the United States of America*,  
677 100, 9383–9387.
- 678 Beck, C. (2009). [Generalised information and entropy measures in physics](#). *Contemporary Physics*, 50,  
679 495–510.
- 680 Canard, E.F., Mouquet, N., Mouillot, D., Stanko, M., Miklisova, D. & Gravel, D. (2014). [Empirical Evaluation](#)  
681 [of Neutral Interactions in Host-Parasite Networks](#). *The American Naturalist*, 183, 468–479.
- 682 Caruso, T., Clemente, G.V., Rillig, M.C. & Garlaschelli, D. (2022). [Fluctuating ecological networks: A](#)  
683 [synthesis of maximum-entropy approaches for pattern detection and process inference](#). *Methods in Ecology*  
684 [and Evolution](#), 13, 2306–2317.
- 685 Cimini, G., Squartini, T., Saracco, F., Garlaschelli, D., Gabrielli, A. & Caldarelli, G. (2019). [The statistical](#)  
686 [physics of real-world networks](#). *Nature Reviews Physics*, 1, 58–71.
- 687 Cirtwill, A.R., Eklof, A., Roslin, T., Wootton, K. & Gravel, D. (2019). [A quantitative framework for](#)  
688 [investigating the reliability of empirical network construction](#). *Methods in Ecology and Evolution*, 10,  
689 902–911.
- 690 Cohen, J.E., Jonsson, T. & Carpenter, S.R. (2003). [Ecological community description using the food web,](#)  
691 [species abundance, and body size](#). *Proceedings of the National Academy of Sciences*, 100, 1781–1786.
- 692 Delmas, E., Besson, M., Brice, M.H., Burkle, L.A., Dalla Riva, G.V., Fortin, M.J., *et al.* (2019). [Analysing](#)  
693 [ecological networks of species interactions](#). *Biological Reviews*.

- 694 Dickinson, J.L., Zuckerberg, B. & Bonter, D.N. (2010). [Citizen Science as an Ecological Research Tool: Challenges and Benefits](#). In: *Annual Review of Ecology, Evolution, and Systematics, Vol 41* (eds. Futuyma, D.J., Shafer, H.B. & Simberloff, D.). Annual Reviews, Palo Alto, pp. 149–172.
- 697 Dunne, J.A., Williams, R.J. & Martinez, N.D. (2002). [Network structure and biodiversity loss in food webs: Robustness increases with connectance](#). *Ecology Letters*, 5, 558–567.
- 699 Dunning, I., Huchette, J. & Lubin, M. (2017). [JuMP: A Modeling Language for Mathematical Optimization](#).  
700 *SIAM Review*, 59, 295–320.
- 701 Fortuna, M.A. & Bascompte, J. (2006). [Habitat loss and the structure of plant-animal mutualistic networks](#).  
702 *Ecology Letters*, 9, 278–283.
- 703 Fortuna, M.A., Stouffer, D.B., Olesen, J.M., Jordano, P., Mouillot, D., Krasnov, B.R., et al. (2010). [Nestedness versus modularity in ecological networks: Two sides of the same coin?](#) *Journal of Animal Ecology*, 79, 811–817.
- 706 Frank, S.A. & Smith, E. (2011). [A simple derivation and classification of common probability distributions based on information symmetry and measurement scale](#). *Journal of Evolutionary Biology*, 24, 469–484.
- 708 Golub, G.H., Hoffman, A. & Stewart, G.W. (1987). [A generalization of the Eckart-Young-Mirsky matrix approximation theorem](#). *Linear Algebra and its Applications*, 88–89, 317–327.
- 710 Gómez, J.M., Perfectti, F. & Jordano, P. (2011). [The Functional Consequences of Mutualistic Network Architecture](#). *PLOS ONE*, 6, e16143.
- 712 Harremoës, P. & Topsøe, F. (2001). [Maximum Entropy Fundamentals](#). *Entropy*, 3, 191–226.
- 713 Harte, J. & Newman, E.A. (2014). [Maximum information entropy: A foundation for ecological theory](#). *Trends in Ecology & Evolution*, 29, 384–389.
- 715 Harte, J., Zillio, T., Conlisk, E. & Smith, A.B. (2008). [Maximum Entropy and the State-Variable Approach to Macroecology](#). *Ecology*, 89, 2700–2711.
- 717 Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M. & Ladle, R.J. (2015). [Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity](#). *Annual Review of Ecology, Evolution, and Systematics*, 46, 523–549.
- 720 Ings, T.C., Montoya, J.M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C.F., et al. (2009). [Review: Ecological networks beyond food webs](#). *Journal of Animal Ecology*, 78, 253–269.

- 722 Jaynes, E.T. (1957a). [Information Theory and Statistical Mechanics](#). *Physical Review*, 106, 620–630.
- 723 Jaynes, E.T. (1957b). [Information Theory and Statistical Mechanics. II](#). *Physical Review*, 108, 171–190.
- 724 Jordano, P. (2016). [Sampling networks of ecological interactions](#). *Functional Ecology*, 30, 1883–1893.
- 725 Khinchin, A.Y. (2013). [Mathematical Foundations of Information Theory](#). Courier Corporation.
- 726 Kirkpatrick, S., Gelatt, C.D. & Vecchi, M.P. (1983). [Optimization by Simulated Annealing](#). *Science*, 220, 671–680.
- 727
- 728 Kullback, S. & Leibler, R.A. (1951). [On Information and Sufficiency](#). *The Annals of Mathematical Statistics*, 22, 79–86.
- 729
- 730 MacDonald, A.A.M., Banville, F. & Poisot, T. (2020). [Revisiting the Links-Species Scaling Relationship in Food Webs](#). *Patterns*, 0.
- 731
- 732 Martyushev, L.M. & Seleznev, V.D. (2006). [Maximum entropy production principle in physics, chemistry and biology](#). *Physics Reports-Review Section of Physics Letters*, 426, 1–45.
- 733
- 734 McCann, K.S. (2011). [Food Webs \(MPB-50\)](#). *Food Webs (MPB-50)*. Princeton University Press.
- 735 Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D. & Alon, U. (2002). [Network motifs: Simple building blocks of complex networks](#). *Science*, 298, 824–827.
- 736
- 737 Park, J. & Newman, M.E.J. (2004). [Statistical mechanics of networks](#). *Physical Review E*, 70, 066117.
- 738
- 739 Pascual, D. of E. and E.B.M. & Dunne, V.P.J.A. (2006). [Ecological Networks: Linking Structure to Dynamics in Food Webs](#). Oxford University Press, USA.
- 740
- 741 Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006). [Maximum entropy modeling of species geographic distributions](#). *Ecological Modelling*, 190, 231–259.
- 742 Pielou, E.C. (1975). Ecological diversity.
- 743 Poisot, T., Baiser, B., Dunne, J.A., Kefi, S., Massol, F., Mouquet, N., *et al.* (2016). [Mangal - making ecological network analysis simple](#). *Ecography*, 39, 384–390.
- 744
- 745 Poisot, T., Bergeron, G., Cazelles, K., Dallas, T., Gravel, D., MacDonald, A., *et al.* (2021). [Global knowledge gaps in species interaction networks data](#). *Journal of Biogeography*, 48, 1552–1563.
- 746
- 747 Poisot, T. & Gravel, D. (2014). [When is an ecological network complex? Connectance drives degree distribution and emerging network properties](#). *PeerJ*, 2, e251.
- 748

- 749 Poisot, T., Stouffer, D.B. & Gravel, D. (2015). [Beyond species: Why ecological interaction networks vary](#)  
750 [through space and time](#). *Oikos*, 124, 243–251.
- 751 Pomeranz, J.P.F., Thompson, R.M., Poisot, T. & Harding, J.S. (2019). [Inferring predator-prey interactions in](#)  
752 [food webs](#). *Methods in Ecology and Evolution*, 10, 356–367.
- 753 Pomeranz, J.P., Thompson, R.M., Poisot, T., Harding, J.S. & Pomeranz, J.P.F. (2018). [Data from: Inferring](#)  
754 [predator-prey interactions in food webs](#).
- 755 Shannon, C.E. (1948). [A mathematical theory of communication](#). *The Bell System Technical Journal*, 27,  
756 379–423.
- 757 Staniczenko, P.P.A., Kopp, J.C. & Allesina, S. (2013). [The ghost of nestedness in ecological networks](#). *Nature*  
758 *Communications*, 4, 1391.
- 759 Stock, M., Poisot, T. & De Baets, B. (2021). [Optimal transportation theory for species interaction networks](#).  
760 *Ecology and Evolution*, 11, 3841–3855.
- 761 Stouffer, D.B. & Bascompte, J. (2011). [Compartmentalization increases food-web persistence](#). *Proceedings of*  
762 *the National Academy of Sciences*, 108, 3648–3652.
- 763 Stouffer, D.B., Camacho, J., Jiang, W. & Nunes Amaral, L.A. (2007). [Evidence for the existence of a robust](#)  
764 [pattern of prey selection in food webs](#). *Proceedings of the Royal Society B-Biological Sciences*, 274,  
765 1931–1940.
- 766 Strydom, T., Catchen, M.D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., et al. (2021a). [A](#)  
767 [roadmap towards predicting species interaction networks \(across space and time\)](#). *Philosophical*  
768 *Transactions of the Royal Society B: Biological Sciences*, 376, 20210063.
- 769 Strydom, T., Dalla Riva, G.V. & Poisot, T. (2021b). [SVD Entropy Reveals the High Complexity of Ecological](#)  
770 [Networks](#). *Frontiers in Ecology and Evolution*, 9, 623141.
- 771 van der Hoorn, P., Lippner, G. & Krioukov, D. (2018). [Sparse Maximum-Entropy Random Graphs with a Given](#)  
772 [Power-Law Degree Distribution](#). *Journal of Statistical Physics*, 173, 806–844.
- 773 Vázquez, D.P. (2005). [Degree distribution in plant-animal mutualistic networks: Forbidden links or random](#)  
774 [interactions?](#) *Oikos*, 108, 421–426.
- 775 Warren II, R.J., Costa, J.T. & Bradford, M.A. (2022). [Seeing shapes in clouds: The fallacy of deriving](#)  
776 [ecological hypotheses from statistical distributions](#). *Oikos*, 2022, e09315.

- 777 Williams, R.J. (2011). [Biology, Methodology or Chance? The Degree Distributions of Bipartite Ecological](#)
- 778 [Networks. PLOS ONE](#), 6, e17645.
- 779 Williams, R.J. & Martinez, N.D. (2000). [Simple rules yield complex food webs. Nature](#), 404, 180–183.
- 780 Williams, R.J. & Martinez, N.D. (2004). [Limits to Trophic Levels and Omnivory in Complex Food Webs: Theory and Data. The American Naturalist](#), 163, 458–468.
- 781

782 **Supporting information**

783 **S1 Fig. Prediction errors of the absolute number of predators  $k_{in}$  and prey  $k_{out}$ .** Species were ordered  
784 according to their total degree in their network. Networks were sorted into different groups based on their total  
785 number of species. In each panel, each dot corresponds to a single species within one of the networks whose  
786 total species count is within the specified range. The predicted joint degree sequences were obtained after  
787 sampling one realization of the joint degree distribution of maximum entropy for each network while keeping  
788 the total number of interactions constant.

789 **S2 Fig. Prediction errors of the relative number of predators  $k_{in}$  and prey  $k_{out}$ .** Species were ordered  
790 according to (a) their out-degree and (b) their in-degree. The predicted joint degree sequences were obtained  
791 after sampling one realization of the joint degree distribution of maximum entropy for each network while  
792 keeping the total number of interactions constant. Due to significant data overlap, all relationships are  
793 represented as 2D histograms. The color bar indicates the number of species that fall within each bin.

794 **S3 Fig. Probability that a species is isolated in its food web.** We derived many degree distributions of  
795 maximum entropy given a range of values of  $S$  and  $L$  and plotted the probability that a species has a degree  $k$  of  
796 0 (log-scale color bar). Species richness varies between 5 and 100 species, by increment of 5 species. For each  
797 level of species richness, the numbers of interactions correspond to all 20-quantiles of the interval between 0  
798 and  $S^2$ . The black line marks the  $S - 1$  minimum number of interactions required to have no isolated species.

799 **S4 Fig. SVD entropy of empirical and predicted food webs.** (a) Distribution of the SVD entropy of empirical  
800 and maximum entropy food webs. Maximum entropy networks were obtained using the type II heuristic MaxEnt  
801 model based on the joint degree sequence. (b) Distribution of z-scores of the SVD entropy of all empirical food  
802 webs. Z-scores were computed using the mean and standard deviation of the distribution of SVD entropy of  
803 MaxEnt food webs (type II heuristic MaxEnt model). The dashed line corresponds to the median z-score.

804 **S5 Fig. Prediction errors of SVD entropy.** Difference in SVD entropy between maximum entropy and  
805 empirical food webs as a function of (a) the number of interactions, (b) connectance, and (c) species richness.  
806 (d) Standardization of the difference in SVD entropy with respect to species richness as a function of species  
807 richness. The exponential decrease in the difference of SVD entropy per species with species richness offers a  
808 complementary perspective supporting the lack of relationship depicted in panel c. Maximum entropy networks  
809 were obtained using the type II heuristic MaxEnt model based on the joint degree sequence.

810 **S6 Fig. Structure of empirical and maximum entropy food webs as a function of species richness.**

811 Maximum entropy networks were obtained using the type II heuristic MaxEnt model based on the joint degree  
812 sequence. (a) Nestedness (estimated using the spectral radius of the adjacency matrix), (b) the maximum  
813 trophic level, (c) the network diameter, and (d) the SVD entropy were measured on these empirical and  
814 maximum entropy food webs and plotted against species richness. Regression lines are plotted in each panel.

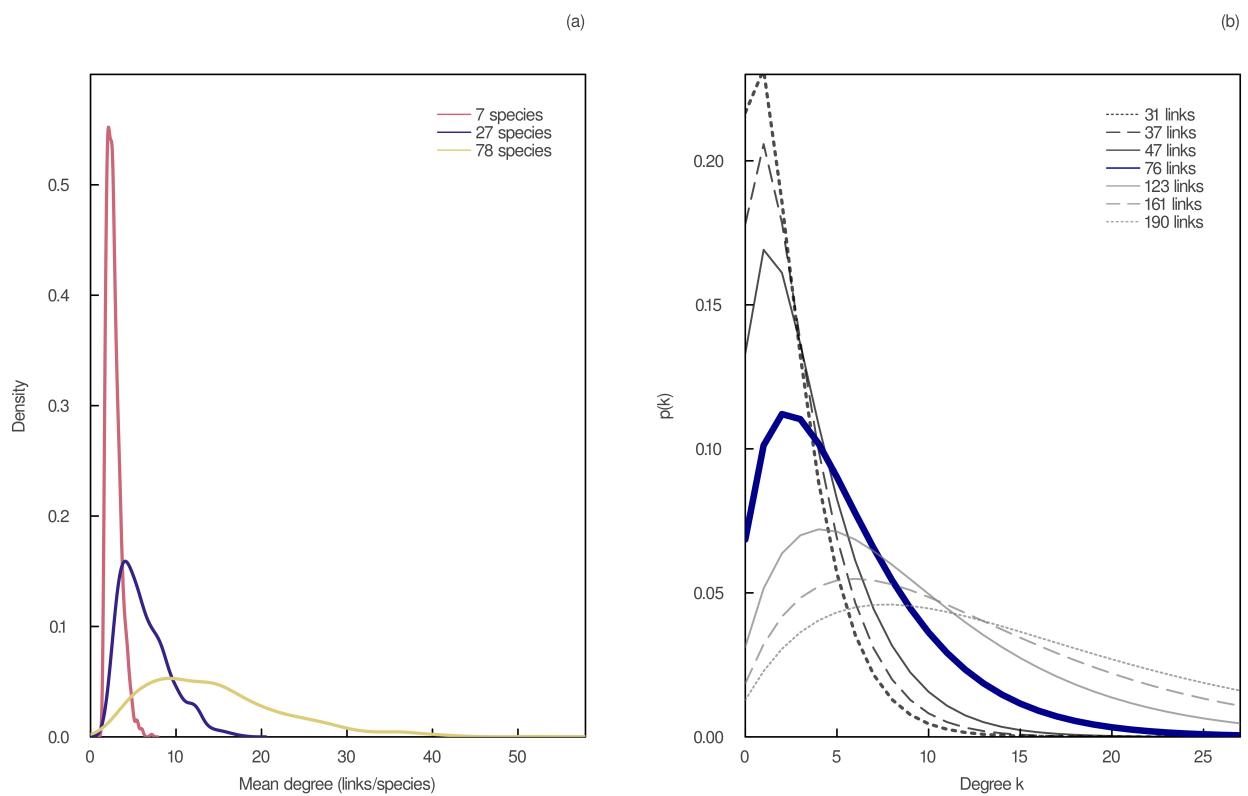


Figure 1: Maximum entropy degree distributions with predicted numbers of interactions.

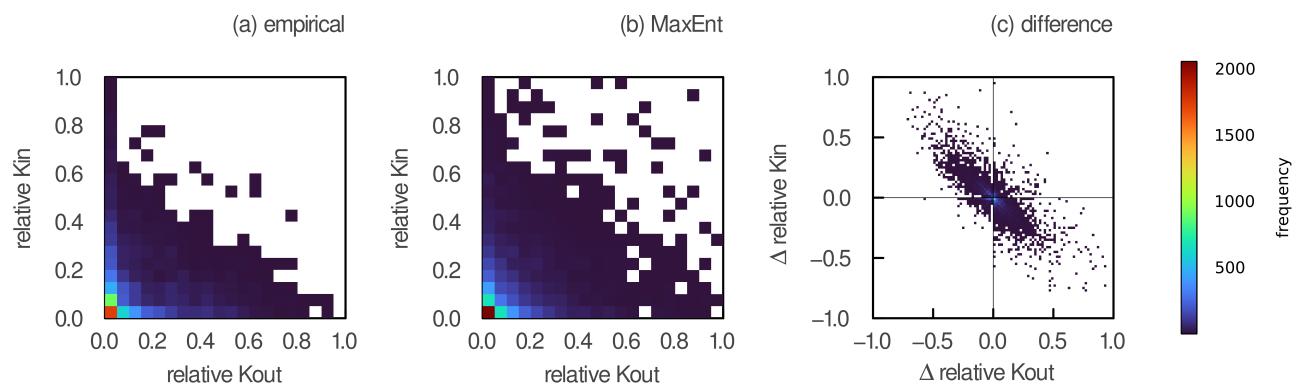


Figure 2: Prediction errors of the relative number of predators and prey.

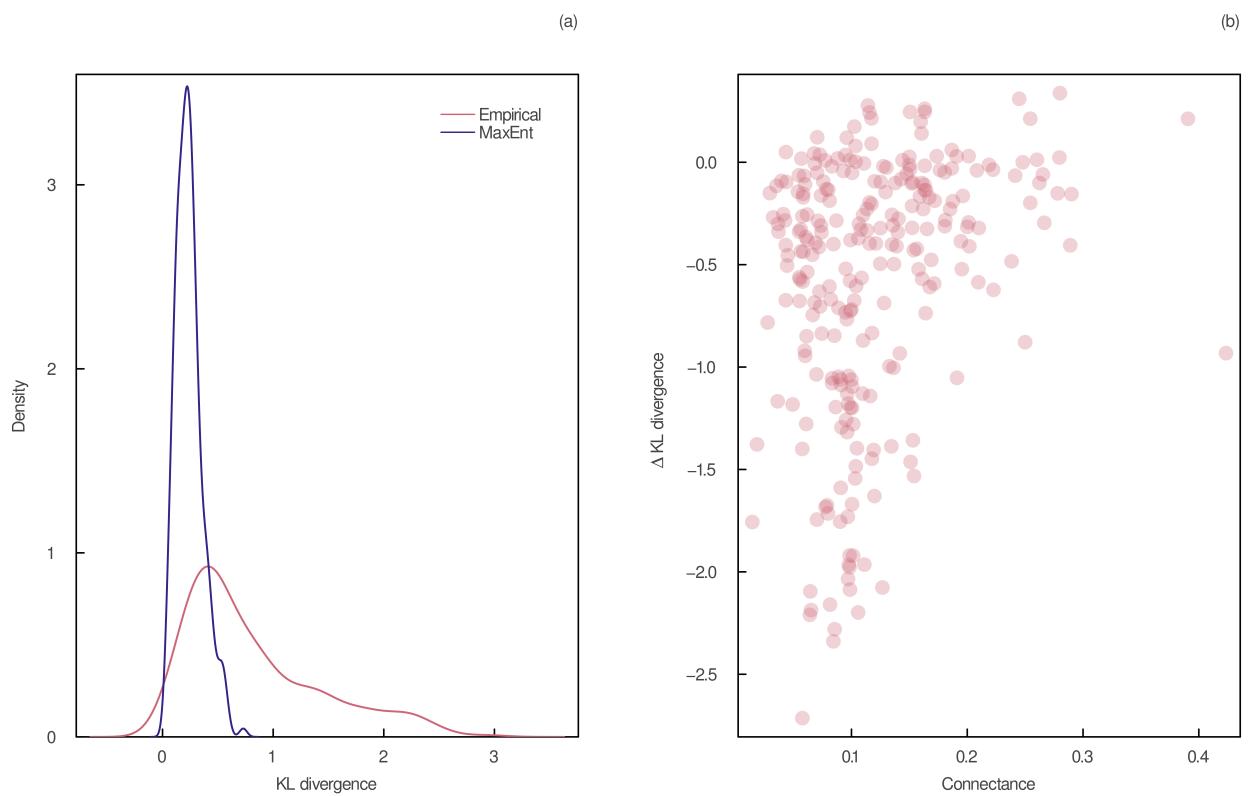


Figure 3: Shape of empirical and predicted joint degree sequences.

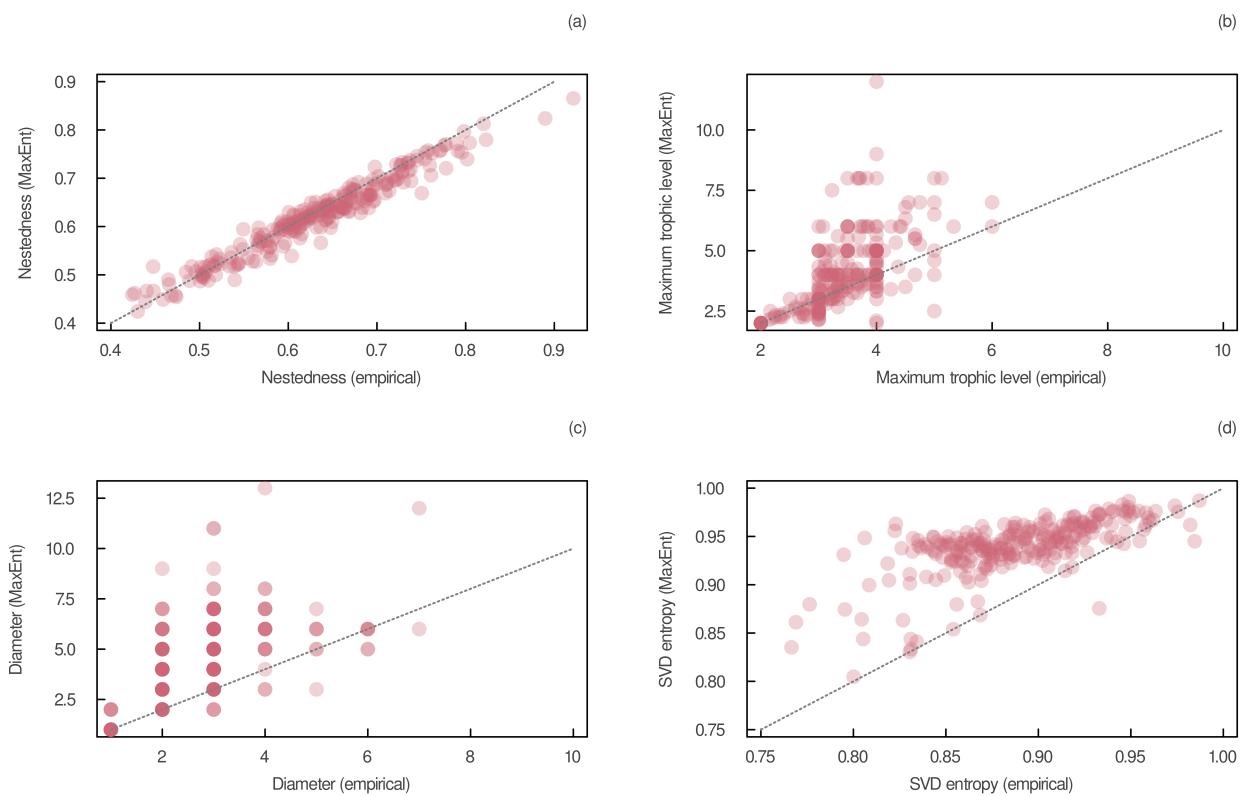


Figure 4: Relationship between the structure of empirical and maximum entropy food webs.

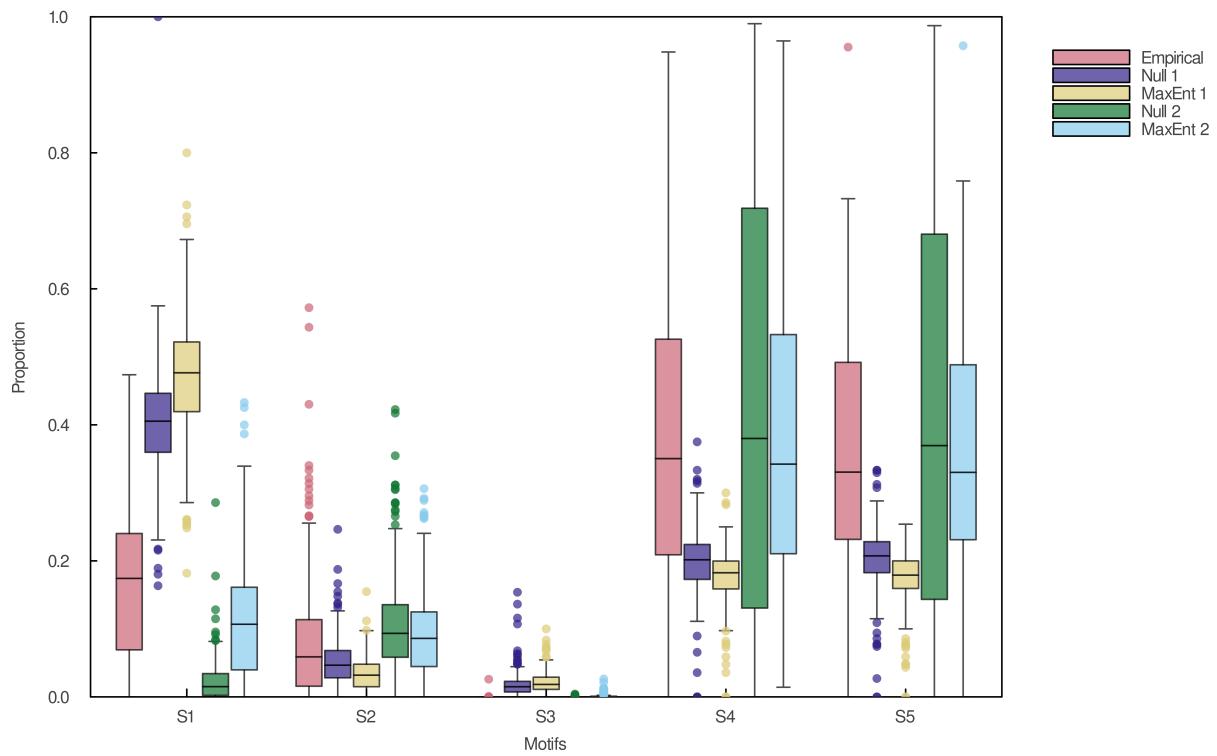


Figure 5: Proportions of single-link three-species motifs in empirical and predicted food webs.

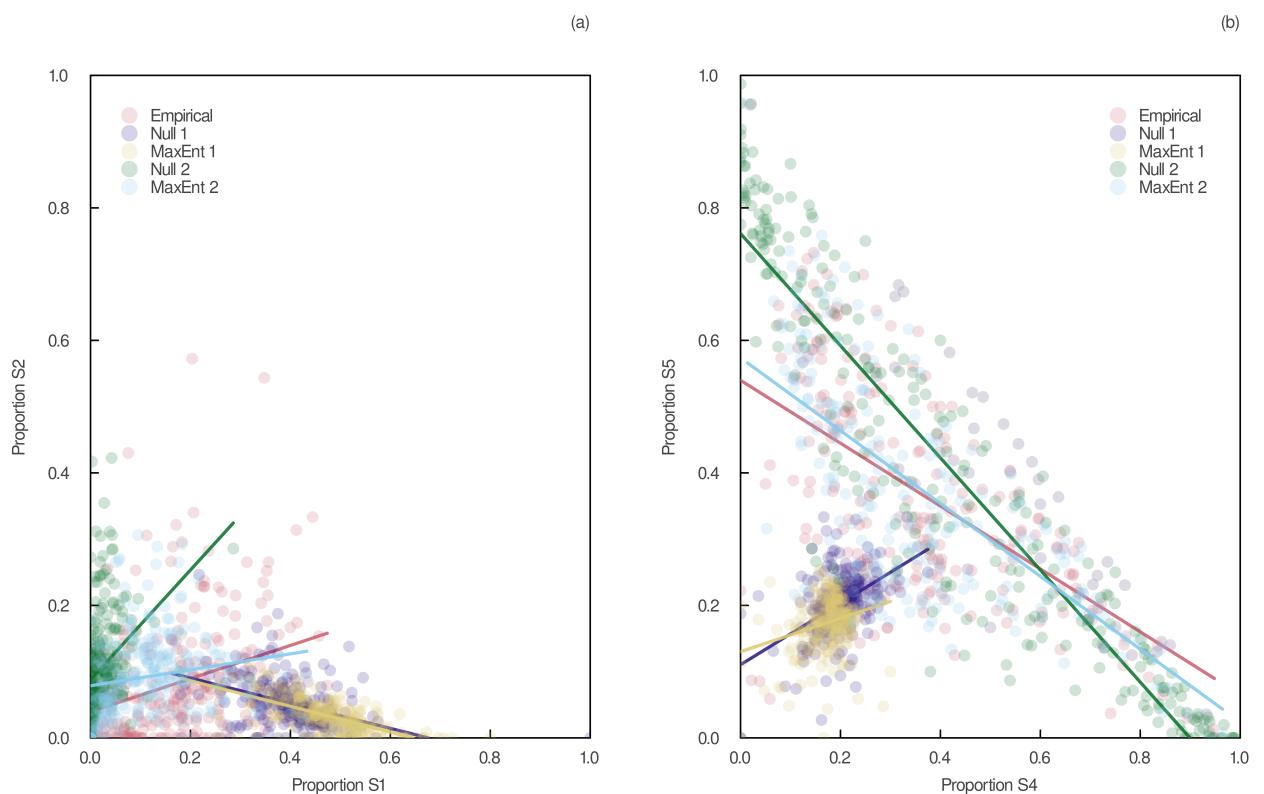


Figure 6: Pairwise relationships between the proportions of single-link three-species motifs in empirical and predicted food webs.