

# Deciphering probabilistic species interaction networks

Francis Banville<sup>1,2,3</sup> Tanya Strydom<sup>1,3</sup> Penelope S. A. Blyth<sup>4</sup> Chris Brimacombe<sup>5</sup> Michael Catchen<sup>3,6</sup>  
Gabriel Dansereau<sup>1,3</sup> Dominique Gravel<sup>2,3</sup> Gracielle Higino<sup>2</sup> Thomas Malpas<sup>4</sup> Hana Mayall<sup>4</sup>  
Kari Norman<sup>1</sup> Timothée Poisot<sup>1,3</sup>

<sup>1</sup> Université de Montréal <sup>2</sup> Université de Sherbrooke <sup>3</sup> Quebec Centre for Biodiversity Science <sup>4</sup> University of Sheffield <sup>5</sup> University of Toronto <sup>6</sup> McGill University

## Correspondance to:

Francis Banville — francis.banville@umontreal.ca

Representing species interactions probabilistically (how likely are they to occur?) as opposed to deterministically (are they occurring?) expresses uncertainties in our knowledge of interactions. The sources of uncertainty captured by interaction probabilities depend on the method used to evaluate them: uncertainty of predictive models, subjective assessment of experts, or empirical measurement of interaction spatiotemporal variability. However, guidelines for the estimation and documentation of probabilistic interaction data are still lacking. This is concerning because our understanding and analysis of interaction probabilities depend on their sometimes elusive definition and uncertainty sources. We review how probabilistic interactions are defined at different spatial scales, from local interactions to regional networks (metawebs), with a strong emphasis on host-parasite and trophic (predatory and herbivory) interactions. These definitions are based on the distinction between the realization of an interaction at a specific time and space (local) and its biological feasibility (regional). Through case studies using host-parasite interactions in Europe, we illustrate how these two network representations differ in their statistical properties, specifically: how local networks and metawebs differ in their spatial and temporal scaling of probabilistic interactions, but not in their taxonomic scaling. We present two approaches to inferring binary interactions from probabilistic ones that account for these differences and show that systematic biases arise when directly inferring local networks from metawebs. Our results underscore the importance of more rigorous descriptions of probabilistic species interaction networks that specify their type of interaction (local or regional) and uncertainty sources.

This work is released by its authors under a CC-BY 4.0 license



Last revision: April 22, 2024

## **1 Introduction**

**2** As we try to navigate global biodiversity change, filling in knowledge gaps about biodiversity becomes  
**3** instrumental to monitoring and mitigating those changes (Abrego *et al.* 2021; Gonzalez & Londoño 2022;  
**4** Hortal *et al.* 2015). However, cataloging species, populations and, in particular, ecological interactions (e.g.,  
**5** predation, parasitism, and pollination) is a substantial challenge (Pascual *et al.* 2006; Polis 1991). There are  
**6** methodological and biological constraints that hinder our ability to document species interactions, inevitably  
**7** leading to uncertainty in our knowledge of interactions. For example, the spatial and temporal uncoupling of  
**8** species (e.g., nocturnal and diurnal species coexisting in the same space with different daily activity timings,  
**9** Jordano 1987) and the large number of rare and cryptic interactions in a community, contribute to these  
**10** knowledge gaps (Jordano 2016).

**11** More generally, several conditions must be satisfied for an interaction to be observed locally. First, both species  
**12** must have overlapping geographic ranges, i.e. they must co-occur within the region of interest (Cazelles *et al.*  
**13** 2016; Morales-Castilla *et al.* 2015). Second, they must have some probability of meeting (Poisot *et al.* 2015).  
**14** Probabilities of interspecific encounters are typically low, especially for rare species with low abundances.  
**15** While species' absolute abundances may impact interaction frequencies (Vázquez *et al.* 2007), encounter  
**16** probabilities are determined by their relative abundances (Canard *et al.* 2012). The probability that species  
**17** meet also depends on their biology, such as their phenology (Olesen *et al.* 2010; Singer & McBride 2012) and  
**18** discoverability (Broom & Ruxton 2005). Finally, when species do come into contact, an interaction occurs only  
**19** if their traits are locally compatible (Poisot *et al.* 2015), including their phenotypes (Bolnick *et al.* 2011; Gravel  
**20** *et al.* 2013; Stouffer *et al.* 2011) and behavior (Choh *et al.* 2012; Pulliam 1974). Interactions may also be  
**21** influenced by a third species (e.g., a more profitable prey species, Golubski & Abrams 2011; Sanders & van  
**22** Veen 2012). Documenting the location and timing of interactions becomes even more difficult when accounting  
**23** for the spatiotemporal variability of ecological interactions (Poisot *et al.* 2012, 2015). Environmental factors,  
**24** such as temperature (Angilletta *et al.* 2004), drought (Woodward *et al.* 2012), climate change (Araujo *et al.*  
**25** 2011; Gilman *et al.* 2010; Woodward *et al.* 2010), and habitat modifications (Tylianakis *et al.* 2007), contribute  
**26** to this spatiotemporal variability by impacting species abundance and traits. Even under favorable  
**27** circumstances, there remains a possibility that the interaction does not occur locally, either due to the intricate  
**28** nature of the system or simply by chance. If it does occur, it might go undetected, particularly if it happens  
**29** infrequently. In this context, it is unsurprising that our knowledge of ecological interactions remains limited

30 (Hortal *et al.* 2015) despite extensive biodiversity data collection (Schmeller *et al.* 2015).

31 Knowing the biological capacity of two species to interact directly (via e.g., trophic interactions) is necessary

32 but not sufficient for inferring their interaction at a specific time and space. The recognition of the intrinsic

33 variability of species interactions has led ecologists to expand their representation of ecological networks to

34 include a probabilistic view of interactions (Dallas *et al.* 2017; Fu *et al.* 2021; Poisot *et al.* 2016). This allows

35 filling in the Eltonian shortfall (Hortal *et al.* 2015) by modeling the probability of occurrence of interactions

36 (e.g., Gravel *et al.* 2019), which can be an important tool for directing efforts and taking action (Carlson *et al.*

37 2021), especially in places where access and resources for research are scarce. The probabilistic representation

38 of interactions has been applied to direct interactions, which are conceptually and mathematically analogous

39 regardless of their biological type (e.g., predation and pollination). This is in contrast with indirect interactions

40 (e.g., interspecific competition), which arise from distinct ecological processes and are often not directly

41 observable (Kéfi *et al.* 2015, 2016).

42 Representing direct interactions probabilistically can capture the spatiotemporal variability of the

43 aforementioned ecological processes and the uncertainty in our knowledge of interactions. Networks of

44 probabilistic interactions, within a Bayesian perspective, express our degree of belief (or confidence) regarding

45 the occurrence or observation of interactions. In contrast, interactions are simply regarded as either occurring or

46 not in networks of deterministic binary interactions. Based on the scale at which they are estimated, interaction

47 probabilities may reflect our level of confidence in whether interactions will be observed, realized locally, or

48 biologically feasible. As an illustration, we could outline a situation in which there is a 50% certainty that an

49 interaction occurs 50% of the time. Our level of confidence should be more definitive (approaching either 0 or

50 1) as we extend our sampling to a broader area and over a longer time period, thereby diminishing the

51 uncertainty of our knowledge of interactions (but not necessarily the estimation of their variability). In the

52 broadest sense, binary interactions are also a type of probabilistic interaction, in which the numerical value of

53 an interaction is restrained to 0 (non-occurring) or 1 (occurring). In networks of probabilistic interactions, only

54 forbidden interactions (i.e., interactions prohibited by biological traits or species absence, Jordano *et al.* 2003;

55 Olesen *et al.* 2010) have a probability value of zero, provided that intraspecific trait variability is considered

56 (Gonzalez-Varo & Traveset 2016).

57 By accounting for the uncertainty of interactions, networks of probabilistic interactions may provide a more

58 realistic portrait of species interactions. However, how the uncertainty of pairwise interactions propagates to

59 network structure (i.e., community-level properties driving the functioning, dynamics, and resilience of

60 ecosystems, McCann 2007; McCann 2011; Proulx *et al.* 2005; Rooney & McCann 2012) remains to be  
61 elucidated. The application and development of computational methods in network ecology, often based on a  
62 probabilistic representation of interactions, can alleviate (and guide) the sampling efforts required to document  
63 species interactions (Strydom *et al.* 2021). For example, statistical models can be used to estimate the  
64 uncertainty of pairwise interactions (Cirtwill *et al.* 2019) and the probability of missing (false negatives) and  
65 spurious (false positives) interactions (Guimerà & Sales-Pardo 2009). Statistical models can generate  
66 predictions of ecological networks without prior knowledge of pairwise interactions. They may do so using  
67 body size (Gravel *et al.* 2013; Petchey *et al.* 2008), phylogeny (Elmasri *et al.* 2020; Strydom *et al.* 2022), or a  
68 combination of niche and neutral processes (Bartomeus *et al.* 2016; Pomeranz *et al.* 2019) for inference.  
69 Topological null models, which generate networks of probabilistic interactions by preserving chosen  
70 characteristics of the adjacency matrix of binary interactions while intentionally omitting others (Bascompte *et*  
71 *al.* 2003; Fortuna & Bascompte 2006), are examples of common probabilistic interaction models. Null models  
72 can be used to produce underlying distributions of network measures for null hypothesis significance testing.  
73 Many measures have been developed to describe the structure (Poisot *et al.* 2016) and diversity (Godsoe *et al.*  
74 2022; Ohlmann *et al.* 2019) of probabilistic interaction networks. These models and measures support the use  
75 of this approach for the study of a wide range of ecological questions, from making better predictions of species  
76 distribution (Cazelles *et al.* 2016) to forecasting the impact of climate change on ecological networks (Gilman  
77 *et al.* 2010).  
78 Yet, a precise definition of probabilistic interactions appears to be lacking, making the estimation and use of  
79 these data more difficult. In this manuscript, we aim to take a step back by outlining different ways in which  
80 probabilistic interactions are defined and used in network ecology. We distinguish two broad categories of  
81 probabilistic interaction networks that necessitate distinct approaches: local networks describing probabilities  
82 of realized interactions, and regional networks (metawebs) describing probabilities of potential interactions. We  
83 highlight the distinctions in the ecological meaning of these two representations and examine some of their  
84 properties and relationships (particularly with space, time, and between each other) through empirical case  
85 studies. Moreover, the lack of clear guidelines on the use of probabilistic interaction data is worrisome, as it  
86 affects both data producers and re-users who generate and manipulate these numbers. This is concerning  
87 because sampling strategies and decisions regarding network construction can affect our understanding of  
88 network properties (Brimacombe *et al.* 2023). There is currently no reporting standard that could guide the  
89 documentation of all types of probabilistic interactions (although see e.g., Salim *et al.* 2022 who discuss data

90 standards for deterministic mutualistic networks). Well-defined reporting standards for probabilistic interactions  
91 would support more adequate manipulation and integration of interaction data from different sources and guard  
92 against possible misinterpretations arising from ambiguous definitions of probabilistic interaction networks.  
93 This documentation should outline the nature (i.e., local or regional) and type (e.g., predatory or pollination) of  
94 the interactions, provide information regarding the taxonomic level, identities, and characteristics (e.g., life  
95 stages) of the individuals involved in an interaction, present the mathematical formulation of probabilities,  
96 including clearly identified conditional variables (e.g., spatial and temporal scales), and describe the methods  
97 and contexts (e.g., location, time, environmental conditions) in which interactions were estimated. Inadequately  
98 documented probabilistic interaction data should be used with caution when analyzing ecological networks.  
99 The broad principles underlying our findings remain relevant and applicable across diverse ecological contexts.

## 100 **Probabilistic representations of interactions**

101 Consider a scenario where an avian predator has just established itself in a northern habitat home to a small  
102 rodent. Suppose these species have never co-occurred before, and as a result, their interaction has not been  
103 previously observed. What is the probability that the rodent is part of the diet of the avian predator, or put  
104 differently, what is the probability that they interact? Answering this question requires some clarification, as  
105 there are multiple ways to interpret and calculate interaction probabilities. We could calculate the probability  
106 that the traits of these species match, i.e. that the avian predator possesses the biological attributes to capture  
107 and consume the rodent. We could also calculate the probability that their traits support an interaction under the  
108 typical environmental conditions of the new habitat. For example, because avian predators hunt by sight,  
109 predation could be possible in the absence of snow but highly improbable when snow is present, as rodents may  
110 use it as a shelter to hide from predators. Finally, we could calculate the probability that the avian predator will  
111 consume the rodent at *that* particular location, for which the spatial and temporal boundaries need to be  
112 specified. The estimation of interaction probabilities, whether through predictive models or prior distributions,  
113 hinges on our understanding of these probabilities and the specific ecological processes we aim to capture.

114 An important aspect to consider when using probabilities of interactions is knowing if they describe potential or  
115 realized interactions, as these two types of interactions have distinct conceptual underpinnings and sources of  
116 uncertainty. A potential (regional) interaction is defined as the biological capacity of two taxa to interact (i.e.,  
117 the probability that they interact if they were to encounter each other and given sufficient time) whereas a

realized (local) interaction is the occurrence or observation of this interaction in a well-defined space and time (i.e., the probability that they interact locally). For two co-occurring taxa and over infinite time, the probability of local interaction is equivalent to the probability of regional (potential) interaction. We use the terms *metaweb* (Dunne 2006) to designate regional networks of potential interactions and *local networks* (Poisot *et al.* 2012) for those of realized interactions. Metawebs are the network analogs of the species pool, where local networks originate from a subset of both species (nodes) and interactions (edges) of the regional metaweb (Saravia *et al.* 2022). Without clear documentation, it can be challenging to know if published probabilistic interaction networks describe local or regional interactions (Tbl. 1 provides examples of studies employing both types of probabilistic interaction networks), or if so-called probabilities are in reality a form of interaction score. When probabilistic regional interactions are used and interpreted incorrectly as local interactions (and conversely), this may generate misleading findings during data analysis. We believe that a better understanding of probabilistic local and regional interactions would prevent interpretation errors (e.g., when studying network-area relationships with metawebs or local networks) and facilitate a more adequate use of interaction data. In the following sections, we delve into these distinctions as we scale up from pairwise interactions to higher-level representations of ecological networks, i.e. local networks and metawebs.

### Pairwise interactions: the building blocks of ecological networks

Local networks and metawebs, like any type of network, are made of nodes and edges that may be represented at different levels of organization. The basic units of ecological networks are individuals that interact with each other (e.g., by predation in food webs, Elton 2001), forming individual-based networks (Melián *et al.* 2011). The aggregation of these individuals into more or less homogeneous groups (e.g., populations, species, families, feeding guilds) allows us to represent nodes at broader taxonomic scales, which affects our interpretation of the properties of these systems (Guimarães 2020; Hemprich-Bennett *et al.* 2021).

Ecologists have traditionally represented interactions (edges) as binary objects that were considered realized after observing at least one individual from group  $i$  interact with at least another individual from group  $j$ . In an adjacency matrix  $B$  of binary interactions, the presence or absence of an interaction  $B_{i \rightarrow j}$  between two taxa can be viewed as the result of a Bernoulli trial  $B_{i \rightarrow j} \sim \text{Bernoulli}(P(B_{i \rightarrow j}))$ , with  $P(B_{i \rightarrow j})$  being the probability of interaction. This interaction probability characterizes our limited ecological knowledge and/or the intrinsic spatiotemporal variability of interactions. It may be estimated through predictive models (e.g., those based on

146 biological traits and species abundances) or expert (prior) knowledge about the interaction. In networks of  
147 probabilistic interactions, the edge values  $P(B_{i \rightarrow j})$  are probabilistic events whose only two possible outcomes  
148 are the presence ( $B_{i \rightarrow j} = 1$ ) or absence ( $B_{i \rightarrow j} = 0$ ) of an interaction between each pair of nodes. Depending on  
149 the type of probabilistic interaction network (local or metaweb), the mathematical formulation and interpretation  
150 of stochastic parameters like  $P(B_{i \rightarrow j})$  can be linked to environmental and biological factors such as species  
151 abundance, traits, area, and time, for example using logistic regression with continuous explanatory variables.

152 Predicting the number of local networks in which the interaction between two given taxa occurs can be achieved  
153 by using a Binomial distribution, assuming a constant interaction probability and independence between  
154 networks (trials). When considering uncertainties around the estimation of  $P(B_{i \rightarrow j})$  or its spatiotemporal  
155 variability, a Beta distribution may be used to represent the relative likelihood of different probability values.  
156 For example, when calculating the probability of interaction between two taxa based on their local abundances,  
157 any spatiotemporal fluctuations in their abundances would introduce variability in the interaction probability at  
158 the local scale. If we take into account the stochasticity of the interaction probability, a Beta-Binomial  
159 distribution can be used to predict the number of networks in which the interaction occurs. Empirically  
160 observing an interaction between two taxa at a given location and time provides important information that can  
161 be used to update previous estimates of  $P(B_{i \rightarrow j})$ , informing us on the biological capacity of both taxa to interact  
162 and the environmental conditions that enabled them to interact locally.

163 Even though binary interaction networks constitute a highly valuable source of ecological information (Pascual  
164 *et al.* 2006), they overlook interaction strengths. Represented in a quantitative adjacency matrix  $W$ , interaction  
165 strengths better describe the energy flows, demographic impacts or frequencies of interactions between nodes  
166 (Berlow *et al.* 2004; Borrett & Scharler 2019), with  $W_{i \rightarrow j}$  being a natural  $\mathbb{N}$  or real  $\mathbb{R}$  number depending on the  
167 measure. For example, they may represent local interaction rates between pairs of taxa (e.g., the flower-visiting  
168 rates of pollinators in a mutualistic network, Herrera 1989). When interaction strengths characterize predation  
169 pressure on prey, they can serve as parameters in a Lotka-Volterra model (e.g., Emmerson & Raffaelli 2004).  
170 The extra amount of ecological information in quantitative networks typically comes at a cost of greater  
171 sampling effort and data volume in predictive models (Strydom *et al.* 2021), which can lead to relatively high  
172 levels of uncertainties when inferring quantitative networks with limited data.

173 As for binary interaction networks, the uncertainty and variability of interaction strengths can be represented  
174 probabilistically. However, the need to estimate the probability distribution of interaction strengths makes the  
175 inference of probabilities more challenging in quantitative networks compared to binary interaction networks.

176 Interaction strengths can follow many probability distributions depending on the method. For instance, they can  
177 follow a Poisson distribution  $W_{i \rightarrow j} \sim \text{Poisson}(\lambda_{i \rightarrow j})$  when predicting the number of interactions between pairs  
178 of nodes, with  $\lambda_{i \rightarrow j}$  being the expected rate at which individuals of taxa  $i$  and  $j$  interact (e.g., the average  
179 number of prey  $j$  consumed by all predators  $i$  in a given time period). The Poisson distribution can also be  
180 0-inflated after initially modeling non-interacting taxa (e.g., Boulangeat *et al.* 2012 employ a 0-inflated model  
181 to analyze species abundance following the modeling of species presence and absence), which constitute the  
182 majority of taxa pairs in most local networks due to their typically high sparseness (Jordano 2016). Because of  
183 the methodological difficulties typically encountered when building deterministic quantitative networks (which  
184 are only partially mitigated by models such as Ecopath, Plagányi & Butterworth 2004), binary interaction  
185 networks, which are easier to sample (Jordano 2016) and predict (Strydom *et al.* 2021), have been more  
186 frequently studied and modeled. Moreover, most published probabilistic interaction networks and methods  
187 describe interactions whose outcome is binary (whether interaction probabilities are regarded as constant or  
188 variable, e.g. represented by a Beta distribution), which underlines the need for better guidelines regarding the  
189 interpretation and manipulation of these types of networks first. For these reasons, our primary focus is  
190 interpreting and using interaction probabilities in Bernoulli distributions, in both local networks and metawebs.

## 191 Local networks: communities interacting in space and time

192 Local networks of probabilistic interactions describe how likely taxa are to interact at a given location and time  
193 period. Local interactions are contingent upon the environmental conditions of the community and the matching  
194 of taxa biological traits. In local networks, edges commonly represent our degree of belief that two taxa interact  
195 in nature, but can also represent the probability of empirically *observing* this interaction (Catchen *et al.* 2023).  
196 We define space as the collection of geographic coordinates  $(x, y, z)$ , with  $(x, y)$  representing longitude and  
197 latitude coordinates, and  $z$  denoting either altitudes or depths. These point coordinates delineate the spatial  
198 boundaries of the system, which may be portrayed as a polyhedron. Ecological interactions may exhibit  
199 variations along latitudinal and altitudinal gradients, as evidenced by changes in hummingbird-plant  
200 interactions (Weinstein & Graham 2017a, b) and mosquito biting rates (e.g., Kulkarni *et al.* 2006) at different  
201 elevations. In contrast, time is defined as the specific time period within which interactions were either observed  
202 or predicted. Even though space and time are continuous variables that should yield probability *densities* of  
203 interactions (i.e., relative likelihoods of interactions occurring at infinitesimal locations and instants in time),

these definitions enable them to be conceptualized as distinct patches and time segments. Treating space and time as discrete dimensions aligns with the common sampling methods of ecological networks and provides probabilities of interactions, which can be obtained by integrating probability densities over space and time. Furthermore, we can quantify both an area  $A$  and a duration  $t$ , which can be readily used in spatiotemporal analyses of ecological networks. When studying network-area relationships (NAR, Galiana *et al.* 2018), we anticipate that local probabilities of interactions scale positively with area and duration because taxa have more opportunities to interact as these dimensions expand.

The probability that two taxa  $i$  and  $j$  interact in a local network  $L_{x,y,z,t}$  (spatial and temporal subscripts hereafter omitted or replaced by the shorter subscript  $k$  for clarity) can be conditioned on many environmental and biological factors. In addition to network area (or volume) and duration, they may be conditioned on taxa co-occurrence  $X_{i,j,k}$ , which is usually a Boolean describing if the geographic distributions of both taxa overlap within the study area. As illustrated in Box 1, co-occurrence may be modeled probabilistically, in which case it may conform to a Bernoulli distribution  $X_{i,j,k} \sim \text{Bernoulli}(P(X_{i,k}, X_{j,k}))$ , where  $X_{i,k}$  and  $X_{j,k}$  are the local occurrences of both taxa. The probability of co-occurrence  $P(X_{i,k}, X_{j,k})$  can be estimated through the application of joint species distribution models (e.g., Pollock *et al.* 2014), potentially taking into account biotic interactions (Staniczenko *et al.* 2017). Given that the probability that two non-co-occurring taxa interact locally is zero (i.e.,  $P(L_{i \rightarrow j}|X_{i,j,k} = 0) = 0$ ), the probability of local interaction can be obtained by multiplying the probability of interaction given co-occurrence with the probability of co-occurrence:

$$P(L_{i \rightarrow j}) = P(L_{i \rightarrow j}|X_{i,j,k} = 1) \times P(X_{i,k} = 1, X_{j,k} = 1). \quad (1)$$

Other important factors that can impact interaction probabilities at the local scale are taxa local abundances  $N_{i,k}$  and  $N_{j,k}$ , which affect encounter probabilities (Canard *et al.* 2012), and local traits distributions  $T_{i,k}$  and  $T_{j,k}$  (e.g., body mass, longevity, and habitat breadth, Caron *et al.* 2024), which determine the ability of individuals to interact after encountering each other (Poisot *et al.* 2015). Moreover, local interactions may be conditional on local environmental factors such as temperature (Angilletta *et al.* 2004), precipitation (Woodward *et al.* 2012), habitat structure (Klecka & Boukal 2014), and presence of other taxa in the network (Kéfi *et al.* 2012; Pilosof *et al.* 2017), as described above. Here, we use the variable  $E_k$  to describe the local environmental context in which interaction probabilities were estimated. For example, in a mesocosm experiment estimating interaction probabilities between predators and prey with and without shelters,  $E_k$  would represent the presence or absence

231 of these shelters. Like co-occurrence,  $E_k$  can also be modeled probabilistically when the stochasticity or  
232 uncertainty of environmental factors is considered.  $E_k$  represents all environmental variables that were taken  
233 into consideration when measuring interaction probabilities; it is a subset of all environmental factors acting on  
234 ecological interactions. Finally, local interaction probabilities may be conditioned on higher-level properties of  
235 the network, which we denote by  $f(L)$ . Many topological null models (i.e., statistical models that randomize  
236 interactions by retaining certain properties of the network while excluding others) provide interaction  
237 probabilities from selected measures of network structure, such as connectance (Fortuna & Bascompte 2006)  
238 and the degree distribution (Bascompte *et al.* 2003).

239 Local interactions must be biologically feasible before occurring at a specific time and space. A local  
240 probability of interaction  $P(L_{i \rightarrow j})$  can be expressed as the product of the probability of local interaction given  
241 that the two taxa can potentially interact  $P(L_{i \rightarrow j}|M_{i \rightarrow j} = 1)$ , which we sometimes denote as  $P(L_{i \rightarrow j}|M_{i \rightarrow j})$  for the  
242 sake of simplicity, with their probability of regional interaction  $P(M_{i \rightarrow j})$ :

$$P(L_{i \rightarrow j}) = P(L_{i \rightarrow j}|M_{i \rightarrow j} = 1) \times P(M_{i \rightarrow j} = 1). \quad (2)$$

243 Low values of  $P(L_{i \rightarrow j}|M_{i \rightarrow j})$  indicate that feasible interactions rarely occur locally, intermediate values around  
244 50% suggest considerable spatiotemporal variability, while high values indicate that regional interactions are  
245 nearly always realized locally. The local probability of interaction between a given pair of taxa is always equal  
246 to or below their probability of regional interaction.

247 The probability that two taxa  $i$  and  $j$  interact in a local network  $L$  can thus be conditioned on their co-occurrence  
248  $X_{i,j,k}$  (or more explicitly on their occurrences  $X_{i,k}$  and  $X_{j,k}$ ), local abundances  $N_{i,k}$  and  $N_{j,k}$ , local traits  
249 distributions  $T_{i,k}$  and  $T_{j,k}$ , local environmental conditions  $E_k$ , network area (or volume)  $A$ , time interval  $t$ ,  
250 network properties  $f(L)$  and biological (or regional) feasibility  $P(M_{i \rightarrow j})$ .

251 Although these variables correspond to distinct ecological inquiries or mechanisms related to ecological  
252 interactions, they may covary with each other, such as the possible dependence of  $X_{i,j,k}$  and  $E_k$  on spatial and  
253 temporal scales. When estimating interaction probabilities using e.g. a generalized linear model with multiple  
254 explanatory variables that might not be independent, it may become important to address collinearity. In such a  
255 case, to mitigate this issue, it may be necessary to use variable selection techniques before fitting the model to  
256 data. The probability that a local interaction is realized is described by the following expression when all these  
257 conditional variables are included:

$$P(L_{i \rightarrow j} | X_{i,k}, X_{j,k}, N_{i,k}, N_{j,k}, T_{i,k}, T_{j,k}, E_k, A, t, f(L), M_{i \rightarrow j}) \quad (3)$$

258 The representation of the local context in which probabilities are estimated and the variables that should be  
 259 taken into consideration depend on the study system, the objectives of the study, and the resources available to  
 260 the researchers. In other words, these variables do not systematically need to be accounted for. For example,  
 261 Gravel *et al.* (2019) analyzed local European trophic networks of willow-galling sawflies and their natural  
 262 enemies, depicting binary interactions, all referenced in space and time, to infer probabilities of local  
 263 interactions between co-occurring species. This was achieved by including temperature and precipitation as  
 264 conditional variables in their models. In Box 2, we show the extent of variation among these local host-parasite  
 265 networks. We do so by measuring their dissimilarity with the regional network (metaweb aggregating all local  
 266 interactions), both in terms of species composition and interactions. We build local probabilistic networks  
 267 following eq. 2, showing that insufficient local variation (high probability of local interaction among potentially  
 268 interacting species) results in an overestimation in both the number of interactions and connectance (i.e., the  
 269 proportion of all of the non-forbidden links that are realized).

270 When accounted for, conditional variables should be clearly described in the documentation of the data  
 271 (Brimacombe *et al.* 2023), preferentially in mathematical terms to avoid any confusion in their interpretation  
 272 and to limit manipulation errors during their re-use. For instance, ecologists should be explicit about their  
 273 consideration of co-occurrence in their estimation of local interaction probabilities. Indeed, it is important to  
 274 specify if probability values are conditional ( $P(L_{i \rightarrow j} | X_{i,j,k} = 1)$ ) or not ( $P(L_{i \rightarrow j})$ ) on co-occurrence, as this can  
 275 change the interpretation of the data. In Tbl. 1, we present examples of studies that used these different  
 276 formulations of probabilistic interactions and conditional variables. We have included the probability of  
 277 empirically observing an interaction that is realized locally  $P(O_{i \rightarrow j} | L_{i \rightarrow j})$  to underscore the distinction between  
 278 local observations and actual realizations of interactions, even though the focus of this manuscript is not on the  
 279 observation of interactions.

**Table 1: Mathematical expressions of probabilistic interactions.** The probability of interaction between two taxa  $i$  and  $j$  is interpreted differently in a local network  $L$  of realized interactions and a metaweb  $M$  of potential interactions (representing the *biological* feasibility of interactions). Each expression emphasizes a different conditional variable, the ellipsis serving as a placeholder for other variables not explicitly stated in the expression. The outcome of each of these probabilistic events, along with common uncertainty sources (often stemming from the model used for estimation), is presented alongside examples of studies that employed them (with specific variables indicated in parentheses, when applicable). The study marked with an asterisk has been conducted on binary interaction networks. Note that we also included the expression for a local network  $O$  of observed interactions and a metaweb  $M^*$  of potential interactions (representing the *ecological* feasibility of interactions) as they were mentioned in the main text, even though they are not the focus of this manuscript.

Expression	Type	Outcome	Variability sources	Reference
$P(L_{i \rightarrow j} X_{i,k}, X_{j,k}, \dots)$	local	realization of the interaction	spatiotemporal	Gravel <i>et al.</i> (2019)
		given taxa co-occurrence	variability	
$P(L_{i \rightarrow j} N_{i,k}, N_{j,k}, \dots)$	local	realization of the interaction	neutral models	Canard <i>et al.</i> (2014)
		given taxa abundances		
$P(L_{i \rightarrow j} T_{i,k}, T_{j,k}, \dots)$	local	realization of the interaction	trait matching	Caron <i>et al.</i> (2024)
		given local traits	models	
$P(L_{i \rightarrow j} E_k, \dots)$	local	realization of the interaction	environmental-	Gravel <i>et al.</i> (2019)
		given local environmental	based models	
		conditions		
$P(L_{i \rightarrow j} A, \dots)$	local	realization of the interaction in	spatial models	Galiana <i>et al.</i> (2018)
		a given area or volume		
$P(L_{i \rightarrow j} t, \dots)$	local	realization of the interaction	temporal models	Weinstein & Graham (2017a)
		during a given time period		
$P(L_{i \rightarrow j} f(L), \dots)$	local	realization of the interaction	topological null	Fortuna & Bascompte (2006) (connectance)
		given network structure	models	
$P(L_{i \rightarrow j} M_{i \rightarrow j}, \dots)$	local	realization of the interaction	spatiotemporal	this study
		given that the taxa can	variability	
		biologically interact		
$P(O_{i \rightarrow j} L_{i \rightarrow j}, \dots)$	local	observation of the interaction	sampling model	Catchen <i>et al.</i> (2023)
		given that it is realized locally		

Expression	Type	Outcome	Variability sources	Reference
$P(M_{i \rightarrow j}   T_i, T_j)$	regional	biological feasibility of the interaction given regional traits (non-forbiddenness)	trait matching models	Strydom <i>et al.</i> (2022)
$P(M_{i \rightarrow j}^*   T_i, T_j, E)$	regional	ecological feasibility of the interaction given regional traits and environmental conditions	trait matching and environmental-based models	this study

When using multiple models to estimate local interaction probabilities, rather than selecting a single model that best fits the data, model averaging may enhance our estimations. In this approach, the weighting is based on the likelihood of each model. Model weights represent the probability that each model is the most suitable for explaining the data, and may be measured using Akaike weights (Burnham & Anderson 2004; Wagenmakers & Farrell 2004). For instance, considering two models  $mod_1$  and  $mod_2$  with respective probabilities (or weights)  $P(mod_1)$  and  $P(mod_2)$ , the average probability of interaction  $P(L_{i \rightarrow j} | \dots)$  can be calculated as follows:

$$P(L_{i \rightarrow j} | \dots) = P(L_{i \rightarrow j} | mod_1, \dots) \times P(mod_1) + P(L_{i \rightarrow j} | mod_2, \dots) \times P(mod_2), \quad (4)$$

where the ellipsis serves as a placeholder for the conditional variables incorporated in these models.

### Box 1: A spatiotemporally explicit model of interactions

Ecologists may resort to predictive models (e.g., generative Bayesian models) to reconstruct local networks across time and space (Strydom *et al.* 2021). Uncertainty in model parameters, model structure, and input data leads to probabilistic interactions. We introduce and develop a simple generative mechanistic model for probabilistic local interactions that takes into consideration their spatiotemporal variability (i.e., a spatiotemporally explicit model of local interactions). This model is not designed for regional interactions, which do not vary spatially and temporally. Rather, it may prove valuable for predicting local interactions across time and space by generating new interaction data following parameter inference.

As indicated by Eq. 1, the probability that two taxa  $i$  and  $j$  interact at a specific location and time  $k$  may be determined by the product of their probability of interaction given co-occurrence and their probability of

co-occurrence. First, their co-occurrence probability can be calculated using their individual (marginal) occurrence probabilities  $P(X_{i,k})$  and  $P(X_{j,k})$ . Given that taxa occurrences are not independent of each other, the joint probability of both taxa occurring together can be calculated by multiplying the probability of one taxon being present by the conditional probability of the other occurring when the first one is present. Alternatively, the joint occurrence probability can be obtained by multiplying the marginal probabilities by the strength of association  $\gamma$  between the occurrences of both taxa (Cazelles *et al.* 2016):

$$P(X_{i,k}, X_{j,k}) = P(X_{i,k}|X_{j,k})P(X_{j,k}) = P(X_{i,k})P(X_{j,k})\gamma. \quad (5)$$

When  $\gamma > 1$ , there is a positive association in the geographic distributions of both taxa, indicating that the presence of one taxon enhances the probability of occurrence of the other. Attractions may be the result of positive interactions (e.g., mutualism) or positive/negative interactions (e.g., antagonism between predators and prey, Cazelles *et al.* 2016). In empirical networks,  $\gamma > 1$  holds for the majority of species pairs (Catchen *et al.* 2023). In contrast, repulsions ( $\gamma < 1$ ) may be caused e.g. by strong interspecific competition (Cazelles *et al.* 2016). We model the co-occurrence  $X_{i,j,k}$  of both taxa as the outcome of a Bernoulli trial

$$X_{i,j,k} \sim \text{Bernoulli}(P(X_{i,k}, X_{j,k})). \quad (6)$$

Next, the probability of interaction given co-occurrence can be made temporally explicit by modeling it as a Poisson process with rate parameter  $\lambda_k$ . This parameter represents the local expected frequency of interaction between the taxa within a defined time interval and can be estimated using prior data on interaction strengths, when accessible. The probability that two co-occurring taxa engage in an interaction during a time period  $t_0$  is given by:

$$P(L_{i \rightarrow j}|X_{i,j,k} = 1) = 1 - e^{-\lambda_k t_0}, \quad (7)$$

which tends toward 1 as  $t_0 \rightarrow \infty$  if  $\lambda > 0$ . In other words, two co-occurring taxa with a nonzero rate of interaction will inevitably interact in a sufficiently long time period. It is important to note that the units of  $\lambda_k$  and  $t_0$  are complementary. For instance, if the duration  $t_0$  is measured in months,  $\lambda_k$  would denote the expected number of interactions per month.

The occurrence of an interaction between  $i$  and  $j$  can be modeled as a Bernoulli trial with a probability of  $P(L_{i \rightarrow j})$ . Consequently, a Bayesian inference model can be built based on the preceding equations to estimate the value of the  $\lambda_k$  and  $\gamma$  parameters and generate novel interaction data:

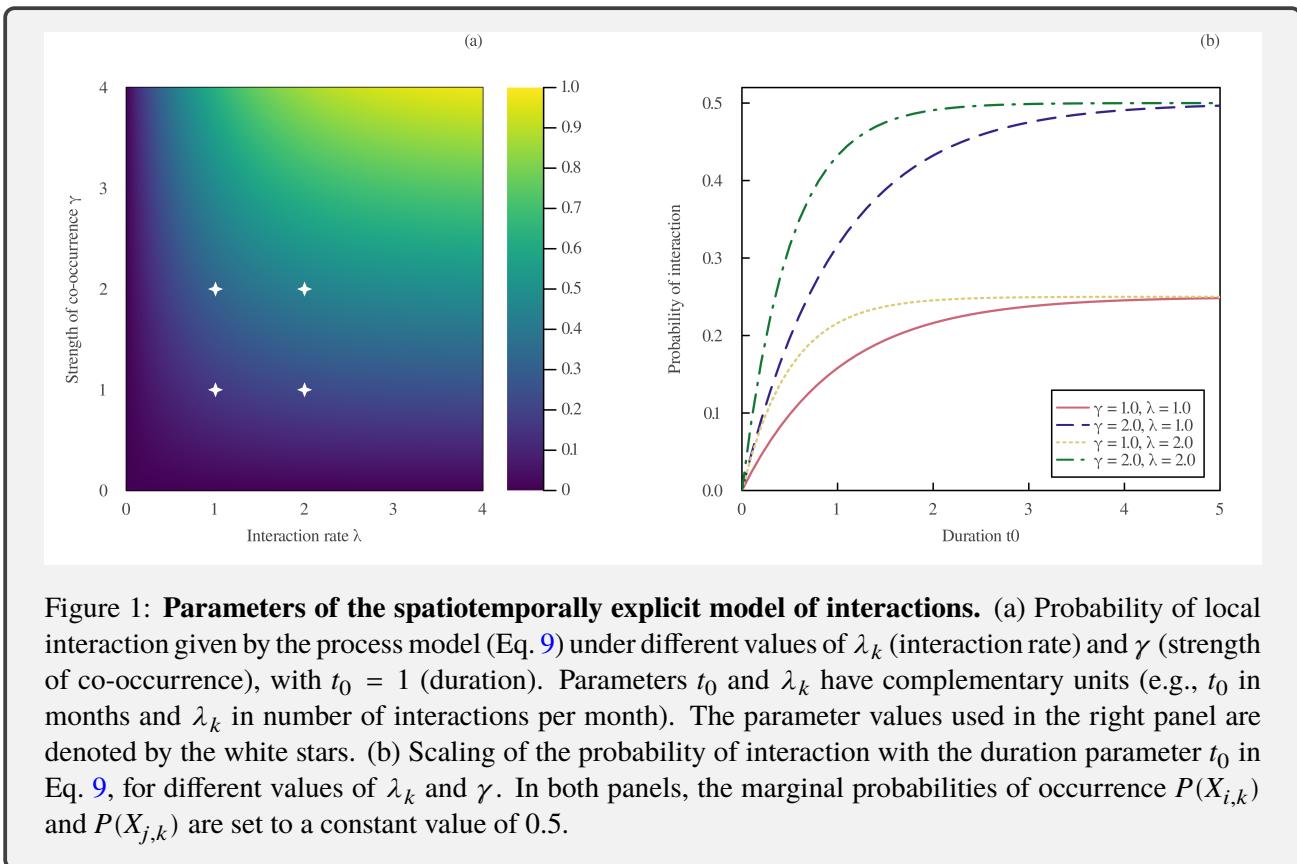
$$L_{i \rightarrow j} \sim \text{Bernoulli}(P(L_{i \rightarrow j})) \quad (8)$$

$$P(L_{i \rightarrow j}) = P(X_{i,k})P(X_{j,k})\gamma(1 - e^{-\lambda_k t_0}) \quad (9)$$

$$\gamma \sim \text{Gamma}(2, 0.5) \quad (10)$$

$$\lambda_k \sim \text{Exponential}(2) \quad (11)$$

In Fig. 1, we show the variation in the probability of interaction under different parameter values. In the right panel, we notice that, irrespective of the interaction rate  $\lambda_k$ , the probability of interaction converges toward an asymptote determined by the probability of co-occurrence  $P(X_{i,j,k})$  (Eq. 5). This model can be customized in different ways, such as by linking  $\lambda_k$  with specific environmental variables or explicitly incorporating observation errors (i.e., the probabilities of false negatives and false positives).



**Figure 1: Parameters of the spatiotemporally explicit model of interactions.** (a) Probability of local interaction given by the process model (Eq. 9) under different values of  $\lambda_k$  (interaction rate) and  $\gamma$  (strength of co-occurrence), with  $t_0 = 1$  (duration). Parameters  $t_0$  and  $\lambda_k$  have complementary units (e.g.,  $t_0$  in months and  $\lambda_k$  in number of interactions per month). The parameter values used in the right panel are denoted by the white stars. (b) Scaling of the probability of interaction with the duration parameter  $t_0$  in Eq. 9, for different values of  $\lambda_k$  and  $\gamma$ . In both panels, the marginal probabilities of occurrence  $P(X_{i,k})$  and  $P(X_{j,k})$  are set to a constant value of 0.5.

290

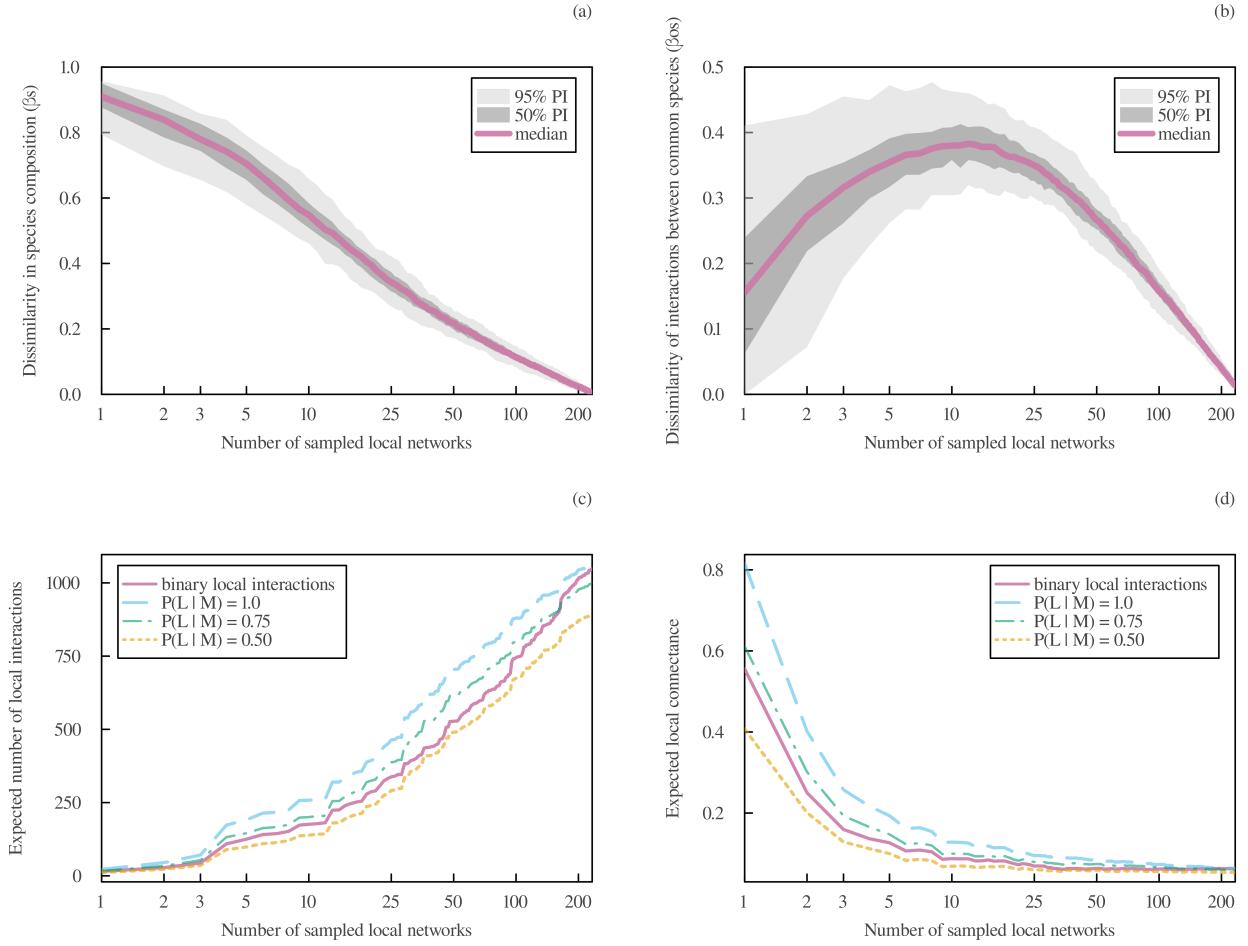
## 291 Box 2: Dissimilarity of local host-parasite networks

292 We use the collection of tripartite host-parasite networks of Kopelke *et al.* (2017), consisting of willows,  
 293 willow-galling sawflies, and their natural enemies, sampled across Europe (233 local networks used). Given its  
 294 replicated networks spanning large spatiotemporal scales, this dataset is well-suited for analyzing network  
 295 variability, both in terms of species composition and interactions. Further details regarding data manipulation  
 296 and network construction can be found in the supplementary material. All code and data to reproduce these  
 297 analyses are available at the Open Science Framework (TBD).

298 In Fig. 2a-b, we show how the dissimilarity between the metaweb of binary interactions and aggregated local  
 299 networks varies with the number of sampled local networks. To do so, we randomly selected one local network  
 300 of binary interactions and sequentially sampled additional networks while aggregating both their species and  
 301 interactions. We compared the metaweb and the aggregated local networks using the dissimilarity in species  
 302 composition ( $\beta_S$ , Fig. 2a) and the dissimilarity of interactions between common species ( $\beta_{OS}$ , Fig. 2b) indices  
 303 (Poisot *et al.* 2012). We repeated this sampling process one hundred times and highlighted the median

304 dissimilarity values across simulations, as well as the 50% and 95% percentile intervals. Both dissimilarity  
305 indices were calculated based on the number of items shared by the two networks ( $c_{LM}$ ) and the number of  
306 items unique to the metaweb ( $u_M$ ) and the aggregated local network ( $u_L$ ). The  $\beta_S$  dissimilarity index uses  
307 species (nodes) as items being compared, while the  $\beta_{OS}$  index assesses dissimilarity based on interactions  
308 between shared species. Both indices were calculated following the  $\beta_W$  index of Whittaker (1960):

$$\beta_W = \frac{c_{LM} + u_L + u_M}{(2c_{LM} + u_L + u_M)/2} - 1. \quad (12)$$



**Figure 2: Network accumulation curves.** (a) Dissimilarity in species composition and (b) dissimilarity of interactions between common species between aggregated local networks and the metaweb of binary host-parasite interactions. Aggregated local networks were obtained by sequentially and randomly selecting a number of local networks of binary interactions and aggregating both their species and interactions. In both panels, the colored line represents the median dissimilarity across simulations and the grey areas cover the 50% and 95% percentile intervals. (c) Scaling of the number of interactions and (d) scaling of connectance with the number of sampled (aggregated) binary and probabilistic local interaction networks. For a better comparison with binary interactions, local networks of probabilistic interactions were derived from a metaweb of probabilistic interactions with a false positive and false negative rate of zero. A specific value of  $P(L_{i \rightarrow j}|M_{i \rightarrow j})$  (the local probability of interaction among potentially interacting species, indices omitted in the figure for simplicity) was used for all local networks within a particular curve. Aggregated local networks of probabilistic interactions were obtained by sequentially and randomly selecting a number of local networks and aggregating both their species and interactions (with the value of  $P(L_{i \rightarrow j}|M_{i \rightarrow j})$  adjusting according to Eq. 13).

309 Expectingly, local networks are highly dissimilar from the metaweb in terms of species composition, especially  
 310 when only a limited number of sites has been sampled. This is because species in the metaweb form the  
 311 regional species pool, and few species occur locally. Moreover, we observe a peak in the dissimilarity of  
 312 interactions between common species at intermediate sampling levels. This suggests that species are collected

313 faster than their interactions. With a limited number of sampled local networks, few regional interactions are  
314 observed locally. Adding more sites brings new species, but not always their interactions. Interactions are more  
315 comprehensively captured when the number of sampled sites is high. Quadratic relationships of network  
316 properties with sampling effort were also observed by McLeod *et al.* (2021).

317 Next, we investigate how the number of local interactions and connectance scale with the number of sampled  
318 (aggregated) local networks of probabilistic or binary interactions (Fig. 2c-d). We built probabilistic local  
319 networks following (**local\_meta?**), and set  $P(M_{i \rightarrow j})$  to 1 when the interaction was observed at least once, and to  
320 0 otherwise (i.e., no false positive or false negative in the metaweb). We used a constant value of  $P(L_{i \rightarrow j} | M_{i \rightarrow j})$   
321 for all interactions.

322 When aggregating local networks of probabilistic interactions, the constancy of the probability of regional  
323 interaction across the entire study area means that any rise in the probability of local interaction is solely  
324 attributable to an increase in  $P(L_{i \rightarrow j} | M_{i \rightarrow j})$ . For example, let  $L_1$  and  $L_2$  be two local networks and  $L_{1,2}$  the  
325 aggregated network. If  $P(L_{1,i \rightarrow j} | M_{i \rightarrow j})$  and  $P(L_{2,i \rightarrow j} | M_{i \rightarrow j})$  are the probabilities that two potentially interacting  
326 taxa interact respectively in  $L_1$  and  $L_2$ , the probability  $P(L_{1,2,i \rightarrow j} | M_{i \rightarrow j})$  that these taxa interact in the aggregated  
327 network  $L_{1,2}$  is obtained by:

$$P(L_{1,2,i \rightarrow j} | M_{i \rightarrow j}) = 1 - [1 - P(L_{1,i \rightarrow j} | M_{i \rightarrow j})] \times [1 - P(L_{2,i \rightarrow j} | M_{i \rightarrow j})], \quad (13)$$

328 assuming independence between the interaction of the two taxa in different networks. This equation represents  
329 the probability that the interaction is realized in either (1) exclusively the local network  $L_1$ , (2) exclusively the  
330 local network  $L_2$  or (3) both, given that the two taxa have the biological capacity to interact.

331 By comparing the scaling relationships observed in local networks of binary (empirical) and probabilistic  
332 (modeled) interactions, Fig. 2c-d shows that high values of  $P(L_{i \rightarrow j} | M_{i \rightarrow j})$  lead to systematic overestimations in  
333 the number of interactions and connectance, especially when  $P(L_{i \rightarrow j} | M_{i \rightarrow j}) = 1$  (corresponding to the scenario  
334 where local probabilities of interactions are equivalent to the probabilities of regional interactions). This  
335 suggests that high values of  $P(L_{i \rightarrow j} | M_{i \rightarrow j})$  do not adequately capture the variability of local interactions.  
336 However, these biases tend to diminish as the number of sampled networks increases, indicating that most  
337 interactions are eventually captured when  $P(L_{i \rightarrow j} | M_{i \rightarrow j})$  is high. In contrast, low values of  $P(L_{i \rightarrow j} | M_{i \rightarrow j})$  lead to  
338 missing interactions, resulting in an underestimation of the number of interactions and connectance. These  
339 results underscore the importance of using the appropriate level of variability when estimating local interaction

340 probabilities.

## 341 Metawebs: regional catalogs of interactions

342 Metawebs (Dunne 2006) are networks of potential interactions over broad spatial, temporal, and taxonomic  
343 scales (e.g., food webs at the continental scale). Potential interactions describe the biological capacity of taxa to  
344 interact, which is typically assessed at the regional scale. Metawebs of probabilistic interactions are particularly  
345 useful in situations where there is uncertainty in the ability of taxa to interact. This uncertainty frequently arises  
346 due to insufficient interaction data, especially for taxa that have not yet been observed to co-occur, and  
347 uncertainties in trait-matching models. As data accumulates, interactions in metawebs should tend towards  
348 binarity, approaching probability values of 0 (repeatedly failing to observe an interaction between two  
349 co-occurring taxa) and 1 (observing an interaction at least once). The extent of sampling effort thus influences  
350 our evaluation of probabilities of potential interactions, as sampling over a larger area or for a longer duration  
351 enables us to capture a greater number of regional interactions (McLeod *et al.* 2021). However, in contrast with  
352 local networks of probabilistic interactions, which describe local interaction stochasticity, regional interactions  
353 are not evaluated for any particular local context. In Box 3, we discuss the differences in spatial and temporal  
354 scaling of regional interactions compared to local interactions. We do so by using the host-parasite networks of  
355 Kopelke *et al.* (2017) as an illustration of spatial scaling. Although *neutrally* forbidden interactions (i.e.,  
356 between rare species, Canard *et al.* 2012) tend to have low probability values in local networks, they may have  
357 higher probabilities in the metaweb (i.e., when species' biological traits can support an interaction if they were  
358 to encounter each other). Likewise, non-co-occurring taxa may have a non-zero probability of interaction in the  
359 metaweb.

360 Potential interactions describe what we refer to as the *biological* feasibility of interactions, which is based solely  
361 on the regional traits distributions  $T_i$  and  $T_j$  of taxa  $i$  and  $j$ , respectively. A probability of potential interaction in  
362 a metaweb  $M$  describing the biological feasibility of interactions may be expressed as:

$$P(M_{i \rightarrow j} | T_i, T_j), \quad (14)$$

363 which, in contrast with local networks, is not conditioned on any spatial, temporal, co-occurrence or  
364 environmental variables (Tbl. 1). Regional traits may differ from local traits  $T_{i,k}$  and  $T_{j,k}$ , which may vary

365 spatially and temporally due to phenotypic plasticity (Berg & Ellers 2010). The taxonomic level at which  
 366 interactions are evaluated influences the distribution of regional and local traits. However, as explained in Box  
 367 4, there is no fundamental difference in the taxonomic scaling of regional and local interactions (i.e., how  
 368 interaction probabilities change with taxonomic level), unlike their spatial and temporal scaling.  
 369 The biological feasibility of interactions expresses our degree of belief that there exists at least one combination  
 370 of phenotypes that could support an interaction if they were to encounter each other, assuming they had infinite  
 371 time to interact. Evaluating this probability is conducted without incorporating the environmental conditions  
 372 under which they encounter each other into the model, as the complement of the probability  $P(F_{i \rightarrow j}|T_i, T_j)$  of  
 373 forbidden interactions based uniquely on biological traits (i.e., the probability that their traits do not support an  
 374 interaction):

$$P(M_{i \rightarrow j}|T_i, T_j) = 1 - P(F_{i \rightarrow j}|T_i, T_j). \quad (15)$$

375 For example, let  $i$  be a western diamondback rattlesnake (*Crotalus atrox*) and  $j$ , a wood lemming (*Myopus*  
 376 *schisticolor*). These two taxa never co-occur, the rattlesnake being adapted to warm regions of North America  
 377 (Castoe *et al.* 2007) and the lemming, to northern habitats of Eurasia (Fedorov *et al.* 2008). As we lack direct  
 378 observations of an interaction between these two species, we have to rely on expert knowledge or trait-matching  
 379 models to estimate their probability of potential interaction. To accurately estimate this probability using  
 380 trait-matching models, it is crucial to ensure that the set of traits considered reflects the overall traits  
 381 distributions of both taxa. We could for instance consider their average body mass and the average phylogenetic  
 382 proximity of lemmings to rattlesnakes' prey. Doing so, we might find a high probability of potential interaction  
 383 based on these traits.

384 The biological feasibility of interactions should not be confused with what we refer to as the *ecological*  
 385 feasibility of interactions. A probability of potential interaction in a metaweb  $M^*$  describing ecological  
 386 feasibility of interactions may be expressed as:

$$P(M_{i \rightarrow j}^*|T_i, T_j, E), \quad (16)$$

387 where  $E$  is the environmental conditions under which potential interactions are evaluated (Tbl. 1). Unlike  $E_k$ ,  
 388 these environmental conditions do not represent conditions occurring at specific locations. Ecological

389 feasibility represents the probability that two taxa interact if they were to encounter each other under given  
390 environmental conditions, assuming they had infinite time to interact. Incorporating environmental conditions  
391 into a trait-matching model may be important when there is high covariance between the environment and traits.  
392 For instance, in our example involving rattlesnakes and lemmings, the probability of potential interaction  
393 between these two species may be low in most environmental conditions. Western diamondback rattlesnakes  
394 may be unactive under low temperatures (Kissner *et al.* 1997), whereas wood lemmings may have low tolerance  
395 to high temperatures (Kausrud *et al.* 2008). The probability that an interaction is ecologically feasible is always  
396 lower than the probability that it is biologically feasible, even across all environmental conditions:

$$\int_E P(M_{i \rightarrow j}^* | T_i, T_j, E) dE \leq P(M_{i \rightarrow j} | T_i, T_j). \quad (17)$$

397 This is because the biological feasibility of an interaction is a prerequisite for its ecological feasibility.  
398 Biological feasibility is necessary but not sufficient for an interaction to be ecologically feasible. Our discussion  
399 of metawebs will focus on the biological feasibility of interactions since most methods developed for inferring  
400 probabilities of regional interactions do not explicitly take into account environmental conditions (e.g., Strydom  
401 *et al.* 2022).  
402 Starting from a selected set of taxa, which are usually distributed within a broad region of interest, metawebs  
403 can be built using different data sources, including literature review (e.g., Maiorano *et al.* 2020), aggregated  
404 interaction data (e.g., Gravel *et al.* 2019; Saravia *et al.* 2022), trait-matching models (e.g., Shaw *et al.* 2024;  
405 Strydom *et al.* 2022), and expert knowledge. Every pair of taxa that have confidently been observed to interact  
406 at least once can be given a probability of 1 (i.e.,  $P(M_{i \rightarrow j}) = 1$ ) since we know that they *can* interact. This  
407 differs from local networks of probabilistic interactions, where interaction events may remain stochastic (i.e.,  
408  $P(L_{i \rightarrow j}) < 1$ ) even after empirically observing interactions due to their spatiotemporal variability. Interactions  
409 that were never observed typically have low probability values in local networks and vary from low to high  
410 values in metawebs, contingent upon taxa traits distributions (reaching 0 for forbidden links).  
411 When incorporating local network data to estimate probabilities of regional interactions, repeatedly failing to  
412 observe an interaction between two co-occurring taxa (i.e.,  $P(O_{i \rightarrow j} = 0)$ ) should decrease the probability that  
413 the interaction is biologically feasible. Using Bayes' theorem, the probability that the interaction is biologically  
414 feasible given that it was never observed locally,  $P(M_{i \rightarrow j} = 1 | O_{i \rightarrow j} = 0, \dots)$ , may be calculated as follows:

$$P(M_{i \rightarrow j} = 1 | O_{i \rightarrow j} = 0, \dots) = \frac{P(O_{i \rightarrow j} = 0 | M_{i \rightarrow j} = 1, \dots) \times P(M_{i \rightarrow j} = 1 | \dots)}{P(O_{i \rightarrow j} = 0 | \dots)}. \quad (18)$$

415 The reduction in the probability of regional interaction after considering that it was never observed locally (i.e.,  
 416  $P(M_{i \rightarrow j} = 1 | O_{i \rightarrow j} = 0, \dots) < P(M_{i \rightarrow j} = 1 | \dots)$ ) occurs because  $P(O_{i \rightarrow j} = 0 | M_{i \rightarrow j} = 1, \dots)$  must be lower than  
 417  $P(O_{i \rightarrow j} = 0, \dots)$ , i.e. there is a higher chance of observing an interaction when we know it is biologically  
 418 feasible.

419 Many observations of interactions are false positives because of observation errors due to taxonomic  
 420 misidentifications and ecological misinterpretations (e.g., due to cryptic species and interactions, Pringle &  
 421 Hutchinson 2020). Likewise, forbidden interactions may be false negatives, e.g. if they have been evaluated  
 422 based on unrepresentative or incomplete traits distributions. Employing Bayesian models proves valuable when  
 423 estimating interaction probabilities in metawebs (e.g., Bartomeus *et al.* 2016; Cirtwill *et al.* 2019). This  
 424 improvement is achieved by updating prior information regarding the feasibility of interactions (e.g., experts'  
 425 prior assessments of interaction probabilities) with empirical data on interactions and traits.

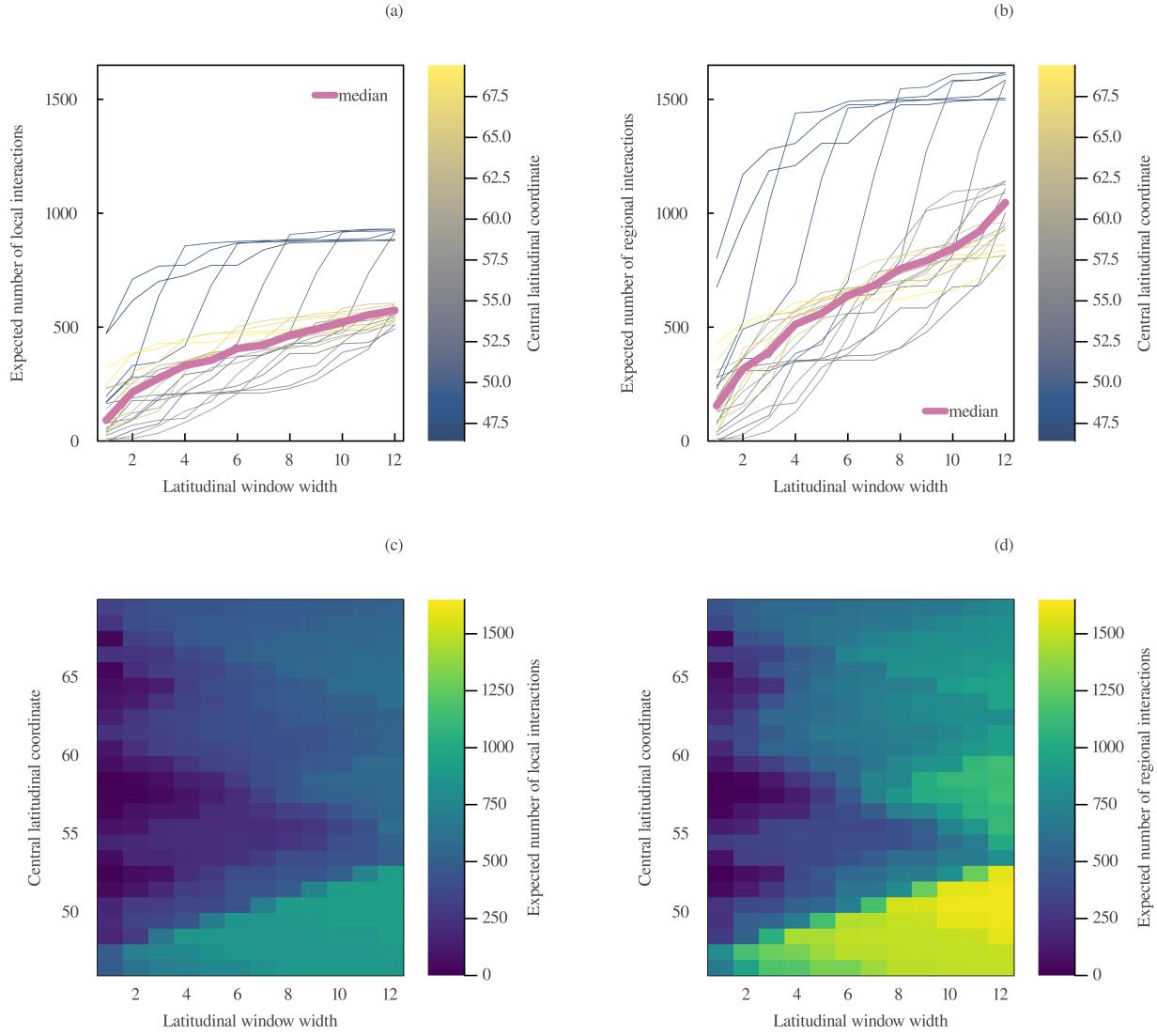
### 426 **Box 3: Spatial and temporal scaling of interactions**

427 Network-Area Relationships document the scaling of network properties (such as modularity and trophic chain  
 428 lengths) with spatial scale (e.g., Galiana *et al.* 2018; Wood *et al.* 2015). The variations in network structure  
 429 across spatial scales may stem from the scaling of species richness (species-area relationships, SARs), the  
 430 number of interactions (Brose *et al.* 2004), and many other higher-level properties of the system (e.g.,  
 431 environmental heterogeneity, Thompson & Townsend 2005) with the sampled area. Additionally, ecological  
 432 processes occurring at distinct spatial scales, such as the spatial variability in local community composition  
 433 resulting from different sequences of extinction and colonization events, can also contribute to this variation  
 434 (Galiana *et al.* 2018). Next, interaction accumulation curves describe the scaling of the number of observed  
 435 interactions with sampling effort (Jordano 2016). Sampling effort, which may correspond to the duration of the  
 436 sampling period used to construct the network, can impact connectance (Bersier *et al.* 1999) and various  
 437 measures of network structure (Banašek-Richter *et al.* 2004; McLeod *et al.* 2021). Apart from sampling effort,  
 438 the temporal scaling of interactions also elucidates how network structure changes with the temporal resolution  
 439 of the network, acknowledging that distinct interactions take place over time, ranging from short-term  
 440 fluctuations of interactions to long-term trends. As local networks of probabilistic interactions may explicitly

441 account for the spatiotemporal variability of interactions, they offer a distinct approach to investigating the  
442 scaling of network structure with space and time, in contrast to networks of binary and quantitative interactions,  
443 by making the stochasticity of interactions the focal point of the modeling process.

444 Local networks and metawebs exhibit distinct relationships with spatial and temporal scales. On one hand,  
445 metawebs of probabilistic interactions, representing biological feasibility, feature regional interactions that do  
446 not scale with space and time. This is because regional interactions depend solely on the biological capacity of  
447 two taxa to interact, regardless of their co-occurrence and local environmental conditions. However,  
448 probabilities of potential interactions may change (tending to become more definitive) upon updating previous  
449 estimates with increased sampling effort, even though they do not vary in a specific direction with the spatial  
450 and temporal extent (boundaries) of the network. The probability of two taxa potentially interacting should  
451 theoretically be the same in all metawebs in which they are present, provided that the data and methods used for  
452 estimation are consistent. For example, if a smaller metaweb  $M_1$  is derived from a larger metaweb  $M_0$  by  
453 selecting the subset of taxa present in the region described by  $M_1$  and retaining all their interactions, their  
454 probabilities of interaction should be identical regardless of scale, i.e.  $P(M_{1,i \rightarrow j}) = P(M_{0,i \rightarrow j})$ . With a larger or  
455 equal number of taxa in  $M_0$  as compared to  $M_1$ , the total number of interactions is expected to be higher or at  
456 least equal in the larger network, even though pairwise probabilities remain identical.

457 On the other hand, local interactions scale both spatially and temporally, given that they have more opportunities  
458 to be realized and observed in larger areas and longer durations. This is attributed to factors such as a higher  
459 number of individuals, greater trait variations, and increased opportunities for encounters, as highlighted by  
460 McLeod *et al.* (2020). For example, if a local network of probabilistic interactions  $L_1$  with an area  $A_1$  is derived  
461 from a larger network  $L_0$  with an area  $A_0$ , and  $A_1$  is entirely nested within  $A_0$ , interaction probabilities should  
462 be lower in the smaller network, i.e.  $P(L_{1,i \rightarrow j}|A_1 < A_0) \leq P(L_{0,i \rightarrow j}|A_0)$ . However, if  $A_1$  and  $A_0$  are disjoint,  
463 interaction probabilities could be higher in the smaller area, contingent upon their environmental and biological  
464 conditions. Likewise, interaction probabilities are expected to be lower in networks with shorter durations when  
465 time intervals are nested. In Fig. 3, we show how the expected number of local host-parasite interactions scales  
466 with area, represented as an expanding latitudinal window, in comparison with regional interactions. Even  
467 though we employed local probabilities of interactions equal to regional interactions for the purpose of  
468 comparison (i.e., using  $P(L_{i \rightarrow j}|M_{i \rightarrow j}) = 1$  here), we notice that the total number of regional interactions scales  
469 more rapidly than local interactions. This is because numerous regional interactions involve species that never  
470 co-occur, and as a result, these interactions are not captured in local networks.



**Figure 3: Spatial scaling of interactions.** Expected number of host-parasite interactions in a network aggregating all (a) local and (b) regional probabilistic interactions within a latitudinal window of a given length. Every dashed curve corresponds to a different expanding window characterized by a different central latitude, with the colored solid line representing the median number of interactions. Heatmaps of the expected number of (c) local and (d) regional interactions found in windows of specified length and central latitudes. Probabilities of regional interactions were obtained with a false positive rate of 5% and a false negative rate of 10%. Local probabilistic interactions were derived from probabilistic regional interactions by setting the value of  $P(L_{i \rightarrow j} | M_{i \rightarrow j})$  (the local probability of interaction among potentially interacting species) to 1, ensuring a conservative comparison between aggregated local networks and metawebs. Aggregated local networks were obtained by aggregating both the species and interactions found within a particular latitudinal window, with the values of  $P(L_{i \rightarrow j} | M_{i \rightarrow j})$  remaining at their maximum value of 1 following Eq. 13.

471 **Box 4: Taxonomic scaling of interactions**

472 Probabilistic interaction networks offer a versatile approach to tackle a broad array of ecological questions,  
473 depending on their level of organization. For instance, the assemblage of interactions across ecological scales  
474 can be explored through species-based networks, while clade-based networks provide insights into  
475 macroevolutionary processes (e.g., Gomez *et al.* 2010). Given that our interpretation of the properties and  
476 dynamics of ecological networks depends on their taxonomic level (Guimarães 2020), investigating the  
477 taxonomic scaling of interactions (i.e., how interaction probabilities change with taxonomic level) emerges as a  
478 promising research avenue. Examining the same system at various taxonomic scales can yield meaningful and  
479 complementary ecological information, and, in our perspective, employing networks of probabilistic  
480 interactions is an effective approach for such analyses.

481 There are no inherent differences in the taxonomic scaling between local and metawebs, as only the nodes are  
482 defined taxonomically. In other words, the probability values of edges in both local (Eq. 3) and metawebs  
483 (Eq. 14) are not conditional on any taxonomic level. The taxonomic scale is tied to the definition of the event  
484 itself (i.e., the interaction between two taxa, defined at the desired taxonomical scale), not to the variables on  
485 which interaction probabilities are conditional. In both types of networks, transitioning to a broader level of  
486 organization (e.g., from a species-level network  $S$  to a genus-level network  $G$ ) can be accomplished directly by  
487 using probabilities from finer scales. For example, in a network with  $n_1$  species from genus  $g_1$  and  $n_2$  species  
488 from genus  $g_2$ , one can compute the probability that at least one species from genus  $g_1$  interacts with at least  
489 one species from genus  $g_2$  (i.e., that there is a non-zero number of species-level interactions or, equivalently,  
490 that the genus-level interaction occurs) as follows:

$$P(G_{g_1 \rightarrow g_2} = 1) = 1 - \prod_{i=1}^{n_1} \prod_{j=1}^{n_2} (1 - P(S_{g_{1i} \rightarrow g_{2j}} = 1)), \quad (19)$$

491 where  $g_{1,i}$  and  $g_{2,j}$  are the species of the corresponding genus and assuming independence between  
492 species-level interactions. If it is known that at least two of these species interact (i.e.,  $P(S_{g_{1,i} \rightarrow g_{2,j}}) = 1$  for at  
493 least one pair of  $(g_{1,i}, g_{2,j})$ ), it implies a probability of genus interaction equal to 1. Canard *et al.* (2012) built a  
494 species-based network following a similar approach, by using simulated interactions between individuals  
495 derived from a neutral model (i.e., a model that assumed ecological equivalence among individuals). In  
496 contrast, a more sophisticated approach is necessary when transitioning from a broader to a finer level of  
497 organization. This is because knowledge of an interaction between two genera does not guarantee that all

498 possible pairwise combinations of their species will also interact. One possible method is to build a finer-scale  
499 network by generating probabilities of interactions through random sampling from a beta distribution,  
500 parameterized by the broader-scale network.

501 Ideally, our biological interpretation of probabilistic interactions should remain consistent across a network  
502 even if it incorporates heterogeneous levels of organization, such as a network whose nodes represent both  
503 species and trophic species (groups of species sharing similar predators and prey). This is common in  
504 ecological networks where taxonomic resolution is typically low (Hemprich-Bennett *et al.* 2021; Vázquez *et al.*  
505 2022). Interaction probabilities at broader taxonomic scales should be based on probabilities of interactions  
506 between individuals, for both local networks and metawebs. For instance, in local individual-based food webs,  
507 the probability that two individuals interact reflects our degree of belief that one individual will consume the  
508 other. Similarly, in local species-based food webs, the probability that two species interact represents our degree  
509 of belief that at least one individual from the predator species consumes at least another individual from the  
510 prey species. Furthermore, in local clade-based food webs, the probability that two clades interact represents  
511 our degree of belief that at least two species from these clades interact with each other or, equivalently, that at  
512 least two individuals from these clades interact with each other. Fundamentally, the taxonomic scaling of  
513 interactions involves aggregating interactions between individuals into larger groups, which may exhibit varying  
514 degrees of homogeneity based on the organisms and the study system. In that regard, taxonomic scaling is  
515 analogous to the spatial and temporal scaling of interactions, as they all represent different ways to aggregate  
516 individuals into broader groups (either spatially, temporally, or taxonomically).

## 517 **Box 5: Sampling for binary interaction networks**

518 The prediction of binary interactions through Bernoulli trials is an important application of probabilistic  
519 interaction networks. This approach proves beneficial for analyzing the structural characteristics of probabilistic  
520 interaction networks, particularly in the absence of specific analytical formulas for measures describing the  
521 structure of the complete network (Poisot *et al.* 2016). By performing independent Bernoulli trials for each  
522 interaction, a network of binary interactions may be generated. A probability distribution of network properties  
523 can be obtained by measuring network structure across multiple randomly generated networks (Poisot *et al.*  
524 2016). This method enables the representation of the variability or uncertainty of network structure, albeit with  
525 possible biases when connectance is low (Chagnon 2015; Poisot & Gravel 2014). Employing this strategy to  
526 generate binary interaction networks under a null model facilitates null hypothesis significance testing, wherein

527 the observed measure is compared against the simulated distribution (e.g., Bascompte *et al.* 2003). Yet,  
 528 inferring network structure through sampling for binary interactions assumes independence among interactions,  
 529 which might not accurately represent reality. Covariation among interactions could exist even if we do not  
 530 explicitly condition interactions on others. For example, an interaction between two taxa could be more  
 531 probable when another interaction occurs. The consequences of this assumption of independence on the  
 532 prediction of network structure have yet to be empirically examined. In Box 5, we argue that predicting binary  
 533 interactions across space using regional interaction probabilities differs from doing so with local interaction  
 534 probabilities. We illustrate both sampling processes using the host-parasite networks of Kopelke *et al.* (2017).  
 535 The random generation of binary interaction networks, from a local network that accounts for the  
 536 spatiotemporal variability of interactions, may effectively capture network structure across space and time,  
 537 facilitating the investigation of ecological hypotheses about interactions at broad spatial and temporal scales.  
 538 There are at least two distinct approaches to sample binary interaction networks from probabilistic interaction  
 539 networks across space, for example, when attempting to predict a binary interaction network for each location  $k$   
 540 within a given region. Both approaches assume independence between interactions. The first approach involves  
 541 performing a singular Bernoulli trial for each pair of taxa based on their regional probability of interaction:

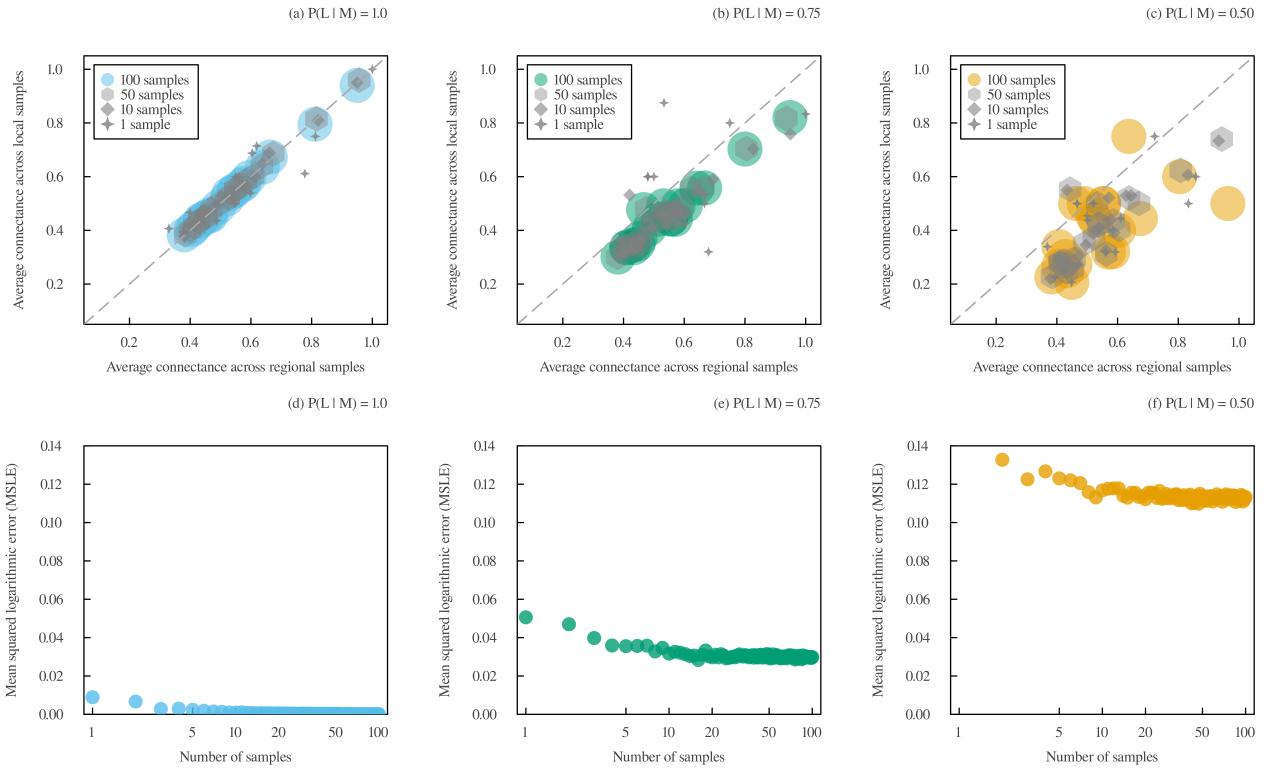
$$M_{i \rightarrow j} \sim \text{Bernoulli}(P(M_{i \rightarrow j})).$$

542 In employing this approach, we predict a single metaweb of binary interactions for each simulation. Every pair  
 543 of taxa predicted to interact in this metaweb will be treated as interacting in all local networks where they  
 544 co-occur, i.e.  $L_{l,i \rightarrow j} = M_{i \rightarrow j}$  when  $X_{i,j,k} = 1$ . This will result in local pairwise interactions without spatial  
 545 variation.  
 546 The second approach is to independently sample each local network of probabilistic interactions:

$$L_{l,i \rightarrow j} \sim \text{Bernoulli}(P(L_{l,i \rightarrow j})).$$

547 This can be achieved by first generating distinct probabilistic interaction networks for each location. These local  
 548 networks of probabilistic interactions may vary in taxa composition and interaction probabilities. Because  
 549 binary interaction networks are sampled independently for each location, this second approach introduces  
 550 spatial variation in binary interactions.

551 In Fig. 4, we compare the average connectance of binary interaction networks resulting from these two sampling  
552 techniques. We drew regional and local interactions from our host-parasite networks of probabilistic  
553 interactions, generating a number of binary interaction network realizations for each site in the dataset. These  
554 two sampling approaches yield different outcomes, particularly for lower values of  $P(L_{i \rightarrow j} | M_{i \rightarrow j})$ , which  
555 represent instances where regional interactions do not consistently manifest locally. Small discrepancies  
556 between these techniques are also apparent when we equate the probability of local interaction to the probability  
557 of regional interaction (i.e., when using  $P(L_{i \rightarrow j} | M_{i \rightarrow j}) = 1.0$  in Eq. 2, Fig. 4a), especially when the number of  
558 simulations of binary interaction networks for each location is low ( $n = 1$ ). As anticipated, we observe that  
559 sampling binary interactions from the metaweb tends to overestimate connectance on average compared to  
560 sampling them from local networks (Fig. 4a-c). Furthermore, we observe an increase in the variability of  
561 connectance when employing a single simulation (Fig. 4, gray markers), which in our opinion is a more tangible  
562 representation of the process leading to the realization of local interactions in nature.



**Figure 4: Connectance of sampled binary interaction networks.** Comparison between the average connectance of binary interaction network samples obtained from the local networks and metawebs of probabilistic interactions. Each dot corresponds to a different site. The local probability of interaction between potentially interacting species was set to three different values: (a)  $P(L_{i \rightarrow j}|M_{i \rightarrow j}) = 1.0$ , (b)  $P(L_{i \rightarrow j}|M_{i \rightarrow j}) = 0.75$ , and (c)  $P(L_{i \rightarrow j}|M_{i \rightarrow j}) = 0.50$ . Grey dots represent the outcome of a single simulation, while colored dots represent the average connectance of each network across 100 simulations. (d) Reduction in the mean squared logarithmic error between the average connectance of binary interaction networks obtained from these two sampling methods as the number of simulations increases, for the same values of  $P(L_{i \rightarrow j}|M_{i \rightarrow j})$  used in panels a-c. Probabilities of regional interactions were obtained with a false positive rate of 5% and a false negative rate of 10%. Metaweb samples were obtained by randomly sampling binary interactions from the probabilistic interaction metaweb, and then propagating this result to all local networks that include the species potentially engaged in the interactions. Local binary interaction networks were generated by independently sampling binary interactions for each local network of probabilistic interactions.

563 The choice of a sampling approach can influence the selection of grid cell size when delineating local  
 564 communities within a broader region of interest. In the first approach, pairwise interactions remain constant  
 565 irrespective of cell size because they are sampled only once from the metaweb. However, in the second  
 566 approach, local interaction probabilities are contingent on network area. For instance, consider the local  
 567 networks  $L_1$  and  $L_2$  with an area of  $\frac{1}{2}A_0$ , both nested within  $A_0$  but disjoint from each other, forming  $L_0$ . If we  
 568 treat  $L_1$  and  $L_2$  as independent, the probability of interaction between taxa  $i$  and  $j$  in  $L_0$  is given by:

$$P(L_{0,i \rightarrow j}) = 1 - [1 - P(L_{1,i \rightarrow j})] \times [1 - P(L_{2,i \rightarrow j})]. \quad (20)$$

569 Due to its larger area, the probability that the two taxa interact in  $L_0$  is equal or greater than in  $L_1$  and  $L_2$ . When  
570 sampling binary interactions from local networks, it is crucial to sample at the same spatial scale for which  
571 probabilities were estimated. Otherwise, interaction probabilities must be adjusted to align with the intended  
572 cell size, preventing systematic biases in predictions.

## 573 Conclusion

574 In this contribution, we underline the importance of network documentation for adequately interpreting and  
575 manipulating probabilistic interaction data. The mathematical representation of probabilities and their statistical  
576 properties depend on the type of interactions (local or regional) and the conditions under which these  
577 interactions were evaluated. We showed that local networks and metawebs of probabilistic interactions differ in  
578 their relationship to spatial and temporal scales (Box 3), with regional interactions remaining consistent across  
579 scales. In contrast with metawebs, local interactions are measured in a specific context (e.g., in a given area,  
580 time, and biological and environmental conditions) and depend on taxa co-occurrence. These differences bring  
581 to light the need to use probabilistic data with caution, for instance when generating network realizations of  
582 binary interactions across space (Box 5). Clear documentation describing the type of interaction and the  
583 variables used in their estimation are required to ensure adequate data manipulation. Sound data practices and  
584 foundations for probabilistic thinking in network ecology facilitate reliable assessments of the spatiotemporal  
585 variability and uncertainty of biotic interactions. Here we identify key research priorities for improving our  
586 understanding of local and regional interactions and their stochasticity.

## 587 Predicting local networks from metawebs

588 Metawebs are a valuable source of ecological information for predicting local networks across time and space.  
589 Local networks of binary interactions can be reconstructed by selecting a subset of taxa and interactions from  
590 the metaweb (Dunne 2006). Metawebs thus contain more interactions than local networks, even though their  
591 connectance is usually much smaller than the one of local networks (Gravel *et al.* 2011). Determining the list of  
592 taxa to select can be achieved empirically (e.g., observed occurrence data for a site) or numerically (e.g., species

593 distribution models). As species composition is arguably easier to sample or predict than pairwise interactions,  
594 the primary challenge lies in deciding which interactions to select from the metaweb. Inferring the structure of  
595 local networks from the metaweb before predicting local pairwise interactions could hold promise (Strydom *et*  
596 *al.* 2021), considering that the structure of local networks is constrained by the metaweb (Saravia *et al.* 2022).

597 Inferring local networks of probabilistic interactions from a metaweb comes with its own set of challenges. For  
598 example, Dansereau *et al.* (2023) inferred spatially-explicit food webs from a metaweb of probabilistic trophic  
599 interactions between Canadian mammals. Their predicted localized food webs are downscaled versions of the  
600 metaweb (i.e., localized metawebs with the same interaction probabilities as those in the regional metaweb). To  
601 infer local networks as defined in this manuscript (i.e., describing local realizations of interactions), local  
602 interaction probabilities must be smaller than regional interaction probabilities. This decrease is due to the  
603 prerequisite that two taxa must initially possess the capacity to interact before engaging in local interactions  
604 (Eq. 2). Inferring local networks from a metaweb by maintaining identical interaction probability values  
605 introduces systematic biases into the predictions, as discussed in Box 1 (unless networks are seen as downscaled  
606 metawebs).

607 As suggested by McLeod *et al.* (2021), although metawebs do not capture the spatiotemporal variability of  
608 interactions, they establish an upper limit for local interactions (similarly for metawebs of probabilistic  
609 interactions, Strydom *et al.* 2023). In other words, the probability that two taxa interact at a specific location  
610 and time is consistently lower or equal to the probability of their regional interaction, regardless of the  
611 conditional variables considered:

$$P(L_{i \rightarrow j} | \dots) \leq P(M_{i \rightarrow j} | T_i, T_j). \quad (21)$$

612 Moreover, the probability that two taxa possess the biological capacity to interact must be higher than the  
613 probability of them interacting at any location and time because they may never co-occur or encounter locally.  
614 Specifically, the cumulative probability of local interaction across all spatial, temporal, and environmental  
615 conditions must be less than the probability of regional interaction, i.e.

$$\int_{E_k} \int_A \int_t P(L_{i \rightarrow j} | E_k, A, t) dt dA dE_k \leq P(M_{i \rightarrow j} | T_i, T_j). \quad (22)$$

616 Estimating more precisely the probability  $P(L_{i \rightarrow j} | M_{i \rightarrow j})$  that two taxa interact locally if they can potentially

617 interact allows for improved predictions of local networks from the metaweb of probabilistic interactions. This  
618 task is challenging due to the variability of this probability across space and time, as well as its variability  
619 across pairwise interactions within a network. Using simple models of  $P(L_{i \rightarrow j} | M_{i \rightarrow j})$ , as shown in Box 1,  
620 represents an initial step toward the overarching objective of reconstructing local networks from metawebs.

## 621 Quantifying and reducing interaction uncertainty

622 While sampling biological communities decreases the uncertainty of interactions by accumulating evidence for  
623 their feasibility and local realization, there is a limit to how much we can reduce uncertainty. In metawebs,  
624 probabilities reflect our limited knowledge of interactions (i.e., our degree of belief that interactions are  
625 feasible), which is expected to improve with a larger volume of data. Regional interactions should become more  
626 definitive (with probabilities approaching 0 or 1) as we investigate various conditions, including different  
627 combinations of species traits.

628 In comparison, although local networks can be seen as random instances of metawebs, their stochasticity cannot  
629 be reduced to the same extent. Local interaction probabilities may represent both their uncertainty and  
630 spatiotemporal variability. Owing to environmental heterogeneity, there will invariably be instances in which an  
631 interaction occurs and others in which it does not, across different times and locations, irrespective of the extent  
632 to which we can improve our knowledge of its biological feasibility and the local conditions that facilitate its  
633 occurrence. When local networks describe probabilities of observing interactions rather than their actual  
634 occurrence, we must also consider observation variability (sampling error) as an additional source of  
635 stochasticity. Every ecological process is stochastic but there is also a possibility that a phenomenon goes  
636 undetected. Quantifying and partitioning this stochasticity will enable us to make more accurate predictions  
637 about ecological interactions at various spatial and temporal scales. This will prove to be of vital importance as  
638 our time to understand nature runs out, especially at locations where the impacts of climate change and habitat  
639 loss hit harder.

## 640 Relaxing the independence assumption

641 Estimating local interaction probabilities independently for each taxa pair and assembling them into a network  
642 of probabilistic interactions comes with limitations. Predicting local networks of binary interactions based on  
643 these interaction probabilities assumes independence among interactions, a condition seldom respected in

644 practice (Golubski & Abrams 2011). Relaxing this assumption is the next logical step in the stochastic  
645 representation of interactions.

646 A more accurate representation of the stochasticity of ecological networks involves creating *probabilistic*  
647 *networks* ( $P(L|...)$  and  $P(M|...)$ ), rather than networks of *probabilistic interactions* ( $P(L_{i \rightarrow j}|...)$  and  $P(M_{i \rightarrow j}|...)$ ).  
648 Probabilistic networks describe the probability that a particular network of binary (or quantitative) interactions  
649 (its whole adjacency matrix) is realized. For example, Young *et al.* (2021) used a Bayesian approach to estimate  
650 the probability of different plant-pollinator network structures derived from imperfect observational data. A  
651 probability distribution of ecological networks may also be derived using the principle of maximum entropy  
652 given structural constrained (e.g., Cimini *et al.* 2019; Park & Newman 2004).

653 Regardless of the method employed, generating probabilistic local networks could lead to more accurate  
654 predictions of local networks of binary interactions by bypassing the independence assumption. Probabilistic  
655 networks could serve as an alternative to null hypothesis significance testing when comparing the structure of a  
656 local network to some random expectations or, as done in Pellissier *et al.* (2018), to the metaweb. These random  
657 expectations are typically derived by performing a series of Bernoulli trials on probabilistic interactions,  
658 assuming independence, to generate a distribution of networks of binary interactions to calculate their structure  
659 (Poisot *et al.* 2016). One could for instance compare the likelihood of an observed network to the one of the  
660 most likely network structure (according to the probabilistic network distribution), thereby directly obtaining a  
661 measure of discrepancy of the empirical network. Generating probabilistic ecological networks represents a  
662 tangible challenge, one that, in the coming years, promises to unlock doors to more advanced and adequate  
663 analyses of ecological networks.

## 664 Acknowledgment

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. This work was  
667 supported by the Institute for Data Valorisation (IVADO) and the Natural Sciences and Engineering Research  
668 Council of Canada (NSERC) Collaborative Research and Training Experience (CREATE) program, through the  
669 Computational Biodiversity Science and Services (BIOS<sup>2</sup>) program. A special thanks to all members of the  
670 Black Holes and Revelations working group (organized by BIOS<sup>2</sup>) for their insightful discussions and valuable  
671 feedback on this manuscript.

672 **References**

- 673 Abrego, N., Roslin, T., Huotari, T., Ji, Y., Schmidt, N.M., Wang, J., *et al.* (2021). Accounting for species  
674 interactions is necessary for predicting how arctic arthropod communities respond to climate change.  
675 *Ecography*, 44, 885–896.
- 676 Angilletta, M.J., Jr., Steury, T.D. & Sears, M.W. (2004). Temperature, Growth Rate, and Body Size in  
677 Ectotherms: Fitting Pieces of a Life-History Puzzle<sup>1</sup>. *Integrative and Comparative Biology*, 44, 498–509.
- 678 Araujo, M.B., Rozenfeld, A., Rahbek, C. & Marquet, P.A. (2011). Using species co-occurrence networks to  
679 assess the impacts of climate change. *Ecography*, 34, 897–908.
- 680 Banašek-Richter, C., Cattin, M.-F. & Bersier, L.-F. (2004). Sampling effects and the robustness of quantitative  
681 and qualitative food-web descriptors. *Journal of Theoretical Biology*, 226, 23–32.
- 682 Bartomeus, I., Gravel, D., Tylianakis, J.M., Aizen, M.A., Dickie, I.A. & Bernard-Verdier, M. (2016). A  
683 common framework for identifying linkage rules across different types of interactions. *Functional Ecology*,  
684 30, 1894–1903.
- 685 Bascompte, J., Jordano, P., Melián, C.J. & Olesen, J.M. (2003). The nested assembly of plant–animal  
686 mutualistic networks. *Proceedings of the National Academy of Sciences*, 100, 9383–9387.
- 687 Berg, M.P. & Ellers, J. (2010). Trait plasticity in species interactions: A driving force of community dynamics.  
688 *Evolutionary Ecology*, 24, 617–629.
- 689 Berlow, E.L., Neutel, A.-M., Cohen, J.E., De Ruiter, P.C., Ebenman, B., Emmerson, M., *et al.* (2004).  
690 Interaction strengths in food webs: Issues and opportunities. *Journal of Animal Ecology*, 73, 585–598.
- 691 Bersier, L.-F., Dixon, P. & Sugihara, G. (1999). Scale-Invariant or Scale-Dependent Behavior of the Link  
692 Density Property in Food Webs: A Matter of Sampling Effort? *The American Naturalist*, 153, 676–682.
- 693 Bolnick, D.I., Amarasekare, P., Araújo, M.S., Bürger, R., Levine, J.M., Novak, M., *et al.* (2011). Why  
694 intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution*, 26, 183–192.
- 695 Borrett, S.R. & Scharler, U.M. (2019). Walk partitions of flow in Ecological Network Analysis: Review and  
696 synthesis of methods and indicators. *Ecological Indicators*, 106, 105451.
- 697 Boulangeat, I., Gravel, D. & Thuiller, W. (2012). Accounting for dispersal and biotic interactions to disentangle  
698 the drivers of species distributions and their abundances. *Ecology Letters*, 15, 584–593.

- 699 Brimacombe, C., Bodner, K., Michalska-Smith, M., Poisot, T. & Fortin, M.-J. (2023). [Shortcomings of reusing](#)  
700 [species interaction networks created by different sets of researchers.](#) *PLOS Biology*, 21, e3002068.
- 701 Broom, M. & Ruxton, G.D. (2005). [You can run—or you can hide: Optimal strategies for cryptic prey against](#)  
702 [pursuit predators.](#) *Behavioral Ecology*, 16, 534–540.
- 703 Brose, U., Ostling, A., Harrison, K. & Martinez, N.D. (2004). [Unified spatial scaling of species and their](#)  
704 [trophic interactions.](#) *Nature*, 428, 167–171.
- 705 Burnham, K.P. & Anderson, D.R. (2004). [Multimodel Inference: Understanding AIC and BIC in Model](#)  
706 [Selection.](#) *Sociological Methods & Research*, 33, 261–304.
- 707 Canard, E.F., Mouquet, N., Mouillot, D., Stanko, M., Miklisova, D. & Gravel, D. (2014). [Empirical Evaluation](#)  
708 [of Neutral Interactions in Host-Parasite Networks.](#) *The American Naturalist*, 183, 468–479.
- 709 Canard, E., Mouquet, N., Marescot, L., Gaston, K.J., Gravel, D. & Mouillot, D. (2012). [Emergence of](#)  
710 [Structural Patterns in Neutral Trophic Networks.](#) *PLOS ONE*, 7, e38295.
- 711 Carlson, C.J., Farrell, M.J., Grange, Z., Han, B.A., Mollentze, N., Phelan, A.L., *et al.* (2021). [The future of](#)  
712 [zoonotic risk prediction.](#) *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376,  
713 20200358.
- 714 Caron, D., Brose, U., Lurgi, M., Blanchet, F.G., Gravel, D. & Pollock, L.J. (2024). [Trait-matching models](#)  
715 [predict pairwise interactions across regions, not food web properties.](#) *Global Ecology and Biogeography*,  
716 33, e13807.
- 717 Castoe, T.A., Spencer, C.L. & Parkinson, C.L. (2007). [Phylogeographic structure and historical demography of](#)  
718 [the western diamondback rattlesnake \(\*Crotalus Atrox\*\): A perspective on North American desert](#)  
719 [biogeography.](#) *Molecular Phylogenetics and Evolution*, 42, 193–212.
- 720 Catchen, M.D., Poisot, T., Pollock, L.J. & Gonzalez, A. (2023). [The missing link: Discerning true from false](#)  
721 [negatives when sampling species interaction networks.](#)
- 722 Cazelles, K., Araujo, M.B., Mouquet, N. & Gravel, D. (2016). [A theory for species co-occurrence in interaction](#)  
723 [networks.](#) *Theoretical Ecology*, 9, 39–48.
- 724 Chagnon, P.-L. (2015). [Characterizing topology of ecological networks along gradients: The limits of metrics'](#)  
725 [standardization.](#) *Ecological Complexity*, 22, 36–39.

- 726 Choh, Y., Ignacio, M., Sabelis, M.W. & Janssen, A. (2012). [Predator-prey role reversals, juvenile experience](#)  
727 [and adult antipredator behaviour](#). *Scientific Reports*, 2, 728.
- 728 Cimini, G., Squartini, T., Saracco, F., Garlaschelli, D., Gabrielli, A. & Caldarelli, G. (2019). [The statistical](#)  
729 [physics of real-world networks](#). *Nature Reviews Physics*, 1, 58–71.
- 730 Cirtwill, A.R., Eklöf, A., Roslin, T., Wootton, K. & Gravel, D. (2019). [A quantitative framework for](#)  
731 [investigating the reliability of empirical network construction](#). *Methods in Ecology and Evolution*, 10,  
732 902–911.
- 733 Dallas, T., Park, A.W. & Drake, J.M. (2017). [Predicting cryptic links in host-parasite networks](#). *PLOS*  
734 [Computational Biology, 13, e1005557.](#)
- 735 Dansereau, G., Barros, C. & Poisot, T. (2023). Spatially explicit predictions of food web structure from regional  
736 level data.
- 737 Dunne, J.A. (2006). The Network Structure of Food Webs. In: *Ecological networks: Linking structure and*  
738 [dynamics](#) (eds. Dunne, J.A. & Pascual, M.). Oxford University Press, pp. 27–86.
- 739 Elmasri, M., Farrell, M.J., Davies, T.J. & Stephens, D.A. (2020). [A hierarchical Bayesian model for predicting](#)  
740 [ecological interactions using scaled evolutionary relationships](#). *The Annals of Applied Statistics*, 14,  
741 221–240.
- 742 Elton, C.S. (2001). *Animal Ecology*. University of Chicago Press, Chicago, IL.
- 743 Emmerson, M.C. & Raffaelli, D. (2004). [Predator–prey body size, interaction strength and the stability of a real](#)  
744 [food web](#). *Journal of Animal Ecology*, 73, 399–409.
- 745 Fedorov, V.B., Goropashnaya, A.V., Boeskorov, G.G. & Cook, J.A. (2008). [Comparative phylogeography and](#)  
746 [demographic history of the wood lemming \(\*Myopus schisticolor\*\): Implications for late Quaternary history](#)  
747 [of the taiga species in Eurasia](#). *Molecular Ecology*, 17, 598–610.
- 748 Fortuna, M.A. & Bascompte, J. (2006). [Habitat loss and the structure of plant–animal mutualistic networks](#).  
749 *Ecology Letters*, 9, 281–286.
- 750 Fu, X., Seo, E., Clarke, J. & Hutchinson, R.A. (2021). [Link Prediction Under Imperfect Detection:](#)  
751 [Collaborative Filtering for Ecological Networks](#). *IEEE Transactions on Knowledge and Data Engineering*,  
752 33, 3117–3128.

- 753 Galiana, N., Lurgi, M., Claramunt-López, B., Fortin, M.-J., Leroux, S., Cazelles, K., *et al.* (2018). [The spatial](#)  
754 [scaling of species interaction networks](#). *Nature Ecology & Evolution*, 2, 782–790.
- 755 Gilman, S.E., Urban, M.C., Tewksbury, J., Gilchrist, G.W. & Holt, R.D. (2010). [A framework for community](#)  
756 [interactions under climate change](#). *Trends in Ecology & Evolution*, 25, 325–331.
- 757 Godsoe, W., Murray, R. & Iritani, R. (2022). [Species interactions and diversity: A unified framework using Hill](#)  
758 [numbers](#). *Oikos*, n/a, e09282.
- 759 Golubski, A.J. & Abrams, P.A. (2011). [Modifying modifiers: What happens when interspecific interactions](#)  
760 [interact?](#) *Journal of Animal Ecology*, 80, 1097–1108.
- 761 Gomez, J.M., Verdu, M. & Perfectti, F. (2010). [Ecological interactions are evolutionarily conserved across the](#)  
762 [entire tree of life](#). *Nature*, 465, 918–U6.
- 763 Gonzalez, A. & Londoño, M.C. (2022). [Monitor biodiversity for action](#). *Science*, 378, 1147–1147.
- 764 Gonzalez-Varo, J.P. & Traveset, A. (2016). [The Labile Limits of Forbidden Interactions](#). *Trends in Ecology &*  
765 *Evolution*, 31, 700–710.
- 766 Gravel, D., Baiser, B., Dunne, J.A., Kopalke, J.-P., Martinez, N.D., Nyman, T., *et al.* (2019). [Bringing Elton](#)  
767 [and Grinnell together: A quantitative framework to represent the biogeography of ecological interaction](#)  
768 [networks](#). *Ecography*, 42, 401–415.
- 769 Gravel, D., Massol, F., Canard, E., Mouillot, D. & Mouquet, N. (2011). [Trophic theory of island biogeography](#).  
770 *Ecology Letters*, 14, 1010–1016.
- 771 Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. (2013). [Inferring food web structure from](#)  
772 [predator–prey body size relationships](#). *Methods in Ecology and Evolution*, 4, 1083–1090.
- 773 Guimarães, P.R. (2020). [The Structure of Ecological Networks Across Levels of Organization](#). *Annual Review*  
774 *of Ecology, Evolution, and Systematics*, 51, 433–460.
- 775 Guimerà, R. & Sales-Pardo, M. (2009). [Missing and spurious interactions and the reconstruction of complex](#)  
776 [networks](#). *Proceedings of the National Academy of Sciences*, 106, 22073–22078.
- 777 Hemprich-Bennett, D.R., Oliveira, H.F.M., Le Comber, S.C., Rossiter, S.J. & Clare, E.L. (2021). [Assessing the](#)  
778 [impact of taxon resolution on network structure](#). *Ecology*, 102, e03256.

- 779 Herrera, C.M. (1989). [Pollinator abundance, morphology, and flower visitation rate: Analysis of the “quantity”](#)  
780 [component in a plant-pollinator system.](#) *Oecologia*, 80, 241–248.
- 781 Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M. & Ladle, R.J. (2015). [Seven Shortfalls](#)  
782 [that Beset Large-Scale Knowledge of Biodiversity.](#) *Annual Review of Ecology, Evolution, and Systematics*,  
783 46, 523–549.
- 784 Jordano, P. (1987). [Patterns of Mutualistic Interactions in Pollination and Seed Dispersal: Connectance,](#)  
785 [Dependence Asymmetries, and Coevolution.](#) *The American Naturalist*, 129, 657–677.
- 786 Jordano, P. (2016). [Sampling networks of ecological interactions.](#) *Functional Ecology*, 30, 1883–1893.
- 787 Jordano, P., Bascompte, J. & Olesen, J.M. (2003). [Invariant properties in coevolutionary networks of](#)  
788 [plant–animal interactions.](#) *Ecology Letters*, 6, 69–81.
- 789 Kausrud, K.L., Mysterud, A., Steen, H., Vik, J.O., Østbye, E., Cazelles, B., *et al.* (2008). [Linking climate](#)  
790 [change to lemming cycles.](#) *Nature*, 456, 93–97.
- 791 Kéfi, S., Berlow, E.L., Wieters, E.A., Joppa, L.N., Wood, S.A., Brose, U., *et al.* (2015). [Network structure](#)  
792 [beyond food webs: Mapping non-trophic and trophic interactions on Chilean rocky shores.](#) *Ecology*, 96,  
793 291–303.
- 794 Kéfi, S., Berlow, E.L., Wieters, E.A., Navarrete, S.A., Petchey, O.L., Wood, S.A., *et al.* (2012). [More than a](#)  
795 [meal... integrating non-feeding interactions into food webs.](#) *Ecology Letters*, 15, 291–300.
- 796 Kéfi, S., Miele, V., Wieters, E.A., Navarrete, S.A. & Berlow, E.L. (2016). [How Structured Is the Entangled](#)  
797 [Bank? The Surprisingly Simple Organization of Multiplex Ecological Networks Leads to Increased](#)  
798 [Persistence and Resilience.](#) *PLOS Biology*, 14, e1002527.
- 799 Kissner, K.J., Forbes, M.R. & Secoy, D.M. (1997). [Rattling Behavior of Prairie Rattlesnakes \(\*Crotalus viridis\*](#)  
800 [viridis, Viperidae\) in Relation to Sex, Reproductive Status, Body Size, and Body Temperature.](#) *Ethology*,  
801 103, 1042–1050.
- 802 Klecka, J. & Boukal, D.S. (2014). [The effect of habitat structure on prey mortality depends on predator and](#)  
803 [prey microhabitat use.](#) *Oecologia*, 176, 183–191.
- 804 Kopelke, J.-P., Nyman, T., Cazelles, K., Gravel, D., Vissault, S. & Roslin, T. (2017). [Food-web structure of](#)  
805 [willow-galling sawflies and their natural enemies across Europe.](#) *Ecology*, 98, 1730–1730.

- 806 Kulkarni, M.A., Kweka, E., Nyale, E., Lyatuu, E., Mosha, F.W., Chandramohan, D., *et al.* (2006).  
807 [Entomological Evaluation of Malaria Vectors at Different Altitudes in Hai District, Northeastern Tanzania.](#)  
808 *Journal of Medical Entomology*, 43, 580–588.
- 809 Maiorano, L., Montemaggiori, A., Ficetola, G.F., O'Connor, L. & Thuiller, W. (2020). [TETRA-EU 1.0: A](#)  
810 [species-level trophic metaweb of European tetrapods](#). *Global Ecology and Biogeography*, 29, 1452–1457.
- 811 McCann, K. (2007). Protecting biostructure. *Nature*, 446, 29–29.
- 812 McCann, K.S. (2011). Food webs (MPB-50). In: *Food Webs (MPB-50)*. Princeton University Press.
- 813 McLeod, A., Leroux, S.J., Gravel, D., Chu, C., Cirtwill, A.R., Fortin, M.-J., *et al.* (2021). [Sampling and](#)  
814 [asymptotic network properties of spatial multi-trophic networks](#). *Oikos*, n/a.
- 815 McLeod, A.M., Leroux, S.J. & Chu, C. (2020). [Effects of species traits, motif profiles, and environment on](#)  
816 [spatial variation in multi-trophic antagonistic networks](#). *Ecosphere*, 11, e03018.
- 817 Melián, C.J., Vilas, C., Baldó, F., González-Ortegón, E., Drake, P. & Williams, R.J. (2011). [Eco-evolutionary](#)  
818 [Dynamics of Individual-Based Food Webs](#). In: *Advances in Ecological Research*, The Role of Body Size in  
819 Multispecies Systems (ed. Belgrano, A.). Academic Press, pp. 225–268.
- 820 Morales-Castilla, I., Matias, M.G., Gravel, D. & Araújo, M.B. (2015). [Inferring biotic interactions from](#)  
821 [proxies](#). *Trends in Ecology & Evolution*, 30, 347–356.
- 822 Ohlmann, M., Miele, V., Dray, S., Chalmandrier, L., O'Connor, L. & Thuiller, W. (2019). [Diversity indices for](#)  
823 [ecological networks: A unifying framework using Hill numbers](#). *Ecology Letters*, 22, 737–747.
- 824 Olesen, J.M., Bascompte, J., Dupont, Y.L., Elberling, H., Rasmussen, C. & Jordano, P. (2010). [Missing and](#)  
825 [forbidden links in mutualistic networks](#). *Proceedings of the Royal Society B: Biological Sciences*, 278,  
826 725–732.
- 827 Park, J. & Newman, M.E.J. (2004). [Statistical mechanics of networks](#). *Physical Review E*, 70, 066117.
- 828 Pascual, M., Dunne, J.A. & Dunne, J.A. (2006). *Ecological Networks: Linking Structure to Dynamics in Food*  
829 *Webs*. Oxford University Press, USA.
- 830 Pellissier, L., Albouy, C., Bascompte, J., Farwig, N., Graham, C., Loreau, M., *et al.* (2018). [Comparing species](#)  
831 [interaction networks along environmental gradients](#). *Biological Reviews*, 93, 785–800.

- 832 Petchey, O.L., Beckerman, A.P., Riede, J.O. & Warren, P.H. (2008). [Size](#), foraging, and food web structure.
- 833 *Proceedings of the National Academy of Sciences*, 105, 4191–4196.
- 834 Pilosof, S., Porter, M.A., Pascual, M. & Kéfi, S. (2017). [The multilayer nature of ecological networks](#). *Nature Ecology & Evolution*, 1, 1–9.
- 835
- 836 Plagányi, É.E. & Butterworth, D.S. (2004). [A critical look at the potential of Ecopath with ecosim to assist in practical fisheries management](#). *African Journal of Marine Science*, 26, 261–287.
- 837
- 838 Poisot, T., Canard, E., Mouillot, D., Mouquet, N. & Gravel, D. (2012). [The dissimilarity of species interaction networks](#). *Ecology Letters*, 15, 1353–1361.
- 839
- 840 Poisot, T., Cirtwill, A.R., Cazelles, K., Gravel, D., Fortin, M.-J. & Stouffer, D.B. (2016). [The structure of probabilistic networks](#). *Methods in Ecology and Evolution*, 7, 303–312.
- 841
- 842 Poisot, T. & Gravel, D. (2014). [When is an ecological network complex? Connectance drives degree distribution and emerging network properties](#). *PeerJ*, 2, e251.
- 843
- 844 Poisot, T., Stouffer, D.B. & Gravel, D. (2015). [Beyond species: Why ecological interaction networks vary through space and time](#). *Oikos*, 124, 243–251.
- 845
- 846 Polis, G.A. (1991). [Complex Trophic Interactions in Deserts: An Empirical Critique of Food-Web Theory](#). *The American Naturalist*, 138, 123–155.
- 847
- 848 Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., *et al.* (2014). [Understanding co-occurrence by modelling species simultaneously with a Joint Species DistributionModel \(JSDM\)](#).
- 849
- 850 *Methods in Ecology and Evolution*, 5, 397–406.
- 851
- 852 Pomeranz, J.P.F., Thompson, R.M., Poisot, T. & Harding, J.S. (2019). [Inferring predator–prey interactions in food webs](#). *Methods in Ecology and Evolution*, 10, 356–367.
- 853
- 854 Pringle, R.M. & Hutchinson, M.C. (2020). [Resolving Food-Web Structure](#). *Annual Review of Ecology, Evolution, and Systematics*, 51, 55–80.
- 855
- 856 Proulx, S.R., Promislow, D.E.L. & Phillips, P.C. (2005). [Network thinking in ecology and evolution](#). *Trends in Ecology & Evolution*, SPECIAL ISSUE: BUMPER BOOK REVIEW, 20, 345–353.
- 857 Pulliam, H.R. (1974). [On the Theory of Optimal Diets](#). *The American Naturalist*, 108, 59–74.

- 858 Rooney, N. & McCann, K.S. (2012). [Integrating food web diversity, structure and stability](#). *Trends in Ecology and Evolution*, 27, 40–46.
- 860 Salim, J.A., Saraiva, A.M., Zermoglio, P.F., Agostini, K., Wolowski, M., Drucker, D.P., *et al.* (2022). [Data standardization of plant–pollinator interactions](#). *GigaScience*, 11, giac043.
- 862 Sanders, D. & van Veen, F.J.F. (2012). [Indirect commensalism promotes persistence of secondary consumer species](#). *Biology Letters*, 8, 960–963.
- 864 Saravia, L.A., Marina, T.I., Kristensen, N.P., De Troch, M. & Momo, F.R. (2022). [Ecological network assembly: How the regional metaweb influences local food webs](#). *Journal of Animal Ecology*, 91, 630–642.
- 866 Schmeller, D.S., Julliard, R., Bellingham, P.J., Böhm, M., Brummitt, N., Chiarucci, A., *et al.* (2015). [Towards a global terrestrial species monitoring program](#). *Journal for Nature Conservation*, 25, 51–57.
- 868 Shaw, J.O., Dunhill, A.M., Beckerman, A.P., Dunne, J.A. & Hull, P.M. (2024). [A framework for reconstructing ancient food webs using functional trait data](#).
- 870 Singer, M.C. & McBride, C.S. (2012). [Geographic mosaics of species' association: A definition and an example driven by plant–insect phenological synchrony](#). *Ecology*, 93, 2658–2673.
- 872 Staniczenko, P.P.A., Sivasubramaniam, P., Suttle, K.B. & Pearson, R.G. (2017). [Linking macroecology and community ecology: Refining predictions of species distributions using biotic interaction networks](#). *Ecology Letters*, 20, 693–707.
- 875 Stouffer, D.B., Rezende, E.L. & Amaral, L.A.N. (2011). [The role of body mass in diet contiguity and food-web structure](#). *Journal of Animal Ecology*, 80, 632–639.
- 877 Strydom, T., Bouskila, S., Banville, F., Barros, C., Caron, D., Farrell, M.J., *et al.* (2022). [Food web reconstruction through phylogenetic transfer of low-rank network representation](#). *Methods in Ecology and Evolution*, 13.
- 880 Strydom, T., Bouskila, S., Banville, F., Barros, C., Caron, D., Farrell, M.J., *et al.* (2023). [Graph embedding and transfer learning can help predict potential species interaction networks despite data limitations](#). *Methods in Ecology and Evolution*, 14, 2917–2930.
- 883 Strydom, T., Catchen, M.D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., *et al.* (2021). [A roadmap towards predicting species interaction networks \(across space and time\)](#). *Philosophical Transactions of the Royal Society B-Biological Sciences*, 376, 20210063.

- 886 Thompson, R.M. & Townsend, C.R. (2005). Food-Web Topology Varies with Spatial Scale in a Patchy  
887 Environment. *Ecology*, 86, 1916–1925.
- 888 Tylianakis, J.M., Tscharntke, T. & Lewis, O.T. (2007). Habitat modification alters the structure of tropical  
889 host-parasitoid food webs. *Nature*, 445, 202–205.
- 890 Vázquez, D.P., Melián, C.J., Williams, N.M., Blüthgen, N., Krasnov, B.R. & Poulin, R. (2007). Species  
891 abundance and asymmetric interaction strength in ecological networks. *Oikos*, 116, 1120–1127.
- 892 Vázquez, D.P., Peralta, G., Cagnolo, L. & Santos, M. (2022). Ecological interaction networks. What we know,  
893 what we don't, and why it matters. *Ecología Austral*, 32, 670–697.
- 894 Wagenmakers, E.-J. & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin &*  
895 *Review*, 11, 192–196.
- 896 Weinstein, B.G. & Graham, C.H. (2017a). On comparing traits and abundance for predicting species  
897 interactions with imperfect detection. *Food Webs*, 11, 17–25.
- 898 Weinstein, B.G. & Graham, C.H. (2017b). Persistent bill and corolla matching despite shifting temporal  
899 resources in tropical hummingbird-plant interactions. *Ecology Letters*, 20, 326–335.
- 900 Whittaker, R.H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological*  
901 *Monographs*, 30, 279–338.
- 902 Wood, S.A., Russell, R., Hanson, D., Williams, R.J. & Dunne, J.A. (2015). Effects of spatial scale of sampling  
903 on food web structure. *Ecology and Evolution*, 5, 3769–3782.
- 904 Woodward, G., Benstead, J.P., Beveridge, O.S., Blanchard, J., Brey, T., Brown, L.E., et al. (2010). Chapter 2 -  
905 Ecological Networks in a Changing Climate. In: *Advances in Ecological Research*, Ecological Networks  
906 (ed. Woodward, G.). Academic Press, pp. 71–138.
- 907 Woodward, G., Brown, L.E., Edwards, F.K., Hudson, L.N., Milner, A.M., Reuman, D.C., et al. (2012). Climate  
908 change impacts in multispecies systems: Drought alters food web size structure in a field experiment.  
909 *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 2990–2997.
- 910 Young, J.-G., Valdovinos, F.S. & Newman, M.E.J. (2021). Reconstruction of plant-pollinator networks from  
911 observational data. *Nature Communications*, 12, 3911.