

# The ecological interpretation of probabilistic 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

The stochastic nature of ecological interactions has led biologists to adopt a probabilistic view of ecological networks. Representing species interactions probabilistically (how likely they are to occur?) as opposed to deterministically (are they occurring?) allows a better assessment of their spatiotemporal variability and accounts for inherent uncertainties in both observations and predictions. However, despite this growing interest in probabilistic networks, guidelines for the estimation and documentation of probabilistic interaction data are still lacking. This is concerning given that our biological understanding of probabilistic interactions and their appropriate analyses depend on the often elusive methods and variables used for their estimation. We review how probabilistic interactions are defined at different spatial scales, from local interactions to regional networks (metawebs), with a strong emphasis on trophic interactions. These definitions are based on the distinction between the observation of an interaction at a specific time and space and its biological feasibility. We show that different network representations have different statistical properties when it comes to common ecological applications. Specifically, we argue that probabilistic local and metawebs differ in their spatial and temporal scaling of interactions, with potential interactions in metawebs remaining constant across spatial and temporal scales. This is in contrast with the taxonomic scaling of interactions, which does not qualitatively differ between both types of networks. We suggest two approaches to inferring deterministic networks from probabilistic webs that account for these differences and argue that systematic biases arise when directly inferring local webs from subsets of metawebs. To support our arguments, we develop different case studies using open empirical data on host-parasite interactions in Europe. Overall, our results underscore the importance of clear metadata for

probabilistic ecological networks, both at the local and regional scales, to inform the appropriate reuse of interaction data.

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



Last revision: *December 12, 2023*

## **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 (Gonzalez & Londoño (2022), Abrego *et al.* (2021)).  
4 However, cataloging species, populations and, in particular, ecological interactions (e.g., predation, parasitism,  
5 and pollination) is a substantial challenge. There are methodological and biological constraints that hinder our  
6 ability to observe all interactions, such as the spatial and temporal uncoupling of species (Jordano (1987)) and  
7 the large number of possible interactions in a community, the vast majority of which are rare (Jordano (2016)).  
8 More generally, a handful of conditions must be satisfied for an interaction to be observed locally. First, both  
9 species must have overlapping geographic ranges, i.e. they must co-occur within the region of interest (Blanchet  
10 *et al.* (2020)). Second, they must have some probability of meeting. Probabilities of interspecific encounters are  
11 typically low, especially for rare species with low relative abundances (Canard *et al.* (2012)). The probability  
12 that species meet each other also depends on their biological characteristics, such as the synchronization of their  
13 phenology (Olesen *et al.* (2010), Singer & McBride (2012)) and their discoverability (e.g., Broom & Ruxton  
14 (2005)). Finally, when species do come into contact, an interaction occurs only if their traits are locally  
15 compatible  
16 (Poisot *et al.* (2015)), including but not limited to their body phenotypes (Bolnick *et al.* (2011), Stouffer *et al.*  
17 (2011), Gravel *et al.* (2013)) and behavioral choices (Pulliam (1974), Choh *et al.* (2012)). Interactions may also  
18 be influenced by the presence or prevalence of a third species (e.g., of a more profitable prey species) (Golubski  
19 & Abrams (2011), Sanders & van Veen (2012)). Documenting the location and timing of interactions becomes  
20 even more difficult when accounting for the spatiotemporal variability of ecological interactions (Poisot *et al.*  
21 (2012), Poisot *et al.* (2015)). Environmental factors, such as temperature (Angilletta *et al.* (2004)), drought  
22 (Woodward *et al.* (2012)), climate change (Gilman *et al.* (2010), Woodward *et al.* (2010), Araujo *et al.* (2011)),  
23 and habitat modifications (Tylianakis *et al.* (2007)), contribute to this spatiotemporal variability of interactions  
24 by impacting species abundance and traits. In this context, it is unsurprising that our knowledge of ecological  
25 interactions remains limited (Hortal *et al.* (2015)) despite extensive biodiversity data collection (Schmeller *et*  
26 *al.* (2015)).  
27 Knowing the biological capacity of two species to interact is necessary but not sufficient for inferring their  
28 interaction at a specific time and space. The recognition of the intrinsic variability of species interactions has  
29 led ecologists to expand their representation of ecological networks (also known as ecological webs) to include

30 a probabilistic view of interactions (Poisot *et al.* (2016), Dallas *et al.* (2017), Fu *et al.* (2021)). This different  
31 perspective allows us to fill in the Eltonian shortfall (Hortal *et al.* (2015)) by modeling the probability of  
32 detecting interactions, which can be an important tool for directing efforts and taking action, especially in  
33 places where access and resources for research are scarce. Representing interactions probabilistically enables us  
34 to capture the spatiotemporal variability of the aforementioned ecological processes and the uncertainties  
35 associated with their measurement. As opposed to binary deterministic webs, in which interactions are regarded  
36 as either occurring or not, probabilistic webs, within a Bayesian framework, express our degree of belief (or  
37 confidence) regarding the occurrence of interactions. Based on the scale at which they are estimated,  
38 probabilistic interactions may reflect our level of confidence in whether interactions will be observed, realized,  
39 or biologically feasible. In the broadest sense, binary networks are also a type of probabilistic network, in which  
40 the numerical value of an interaction is restrained to 0 (non-occurring) or 1 (occurring). In probabilistic webs,  
41 only forbidden interactions (i.e., interactions prohibited by biological traits, Jordano *et al.* (2003), Olesen *et al.*  
42 (2010)) have a probability value of zero, provided that intraspecific trait variability is considered  
43 (Gonzalez-Varo & Traveset (2016)).

44 By accounting for the uncertainty of interactions, probabilistic webs provide a more realistic portrait of species  
45 interactions and network structure (i.e. community-level properties), which is a major driver of the functioning,  
46 dynamics, and resilience of ecosystems worldwide (Proulx *et al.* (2005), McCann (2007), McCann (2011),  
47 Rooney & McCann (2012)). Moreover, the application and development of computational methods in network  
48 ecology, often based on a probabilistic representation of interactions, can alleviate (and guide) the sampling  
49 efforts required to document species interactions (Strydom *et al.* (2021)). For example, statistical models can be  
50 used to estimate the uncertainty of pairwise interactions (Cirtwill *et al.* (2019)) and the probability of missing  
51 (false negatives) and spurious (false positives) interactions (Guimerà & Sales-Pardo (2009)). Considering the  
52 high rate of false negatives in species interaction data due to the difficulty of witnessing rare interactions  
53 (Catchen *et al.* (2023)), these models could inform the identification of priority sampling locations of ecological  
54 webs where data collection would yield the most valuable information, thereby reducing errors. Optimization  
55 models for sampling locations have mostly found applications in biological systems that are not networks, such  
56 as identifying priority sampling sites for disease hotspots (Andrade-Pacheco *et al.* (2020)), but there is  
57 substantial promise in applying them to probabilistic ecological interactions. Statistical models can also be used  
58 to generate predictions of ecological webs without prior knowledge of pairwise interactions, for instance using  
59 body size (Petchey *et al.* (2008), Gravel *et al.* (2013)), phylogeny (Elmasri *et al.* (2020), Strydom *et al.* (2022)),

60 or a combination of niche and neutral processes (Bartomeus *et al.* (2016), Pomeranz *et al.* (2019)) for inference.  
61 Topological null models, which generate probabilistic networks by preserving chosen characteristics of the  
62 binary adjacency matrix while intentionally omitting others (Bascompte *et al.* (2003), Fortuna & Bascompte  
63 (2006)), serve as other examples of common probabilistic network models. Null models can be used to produce  
64 underlying distributions of network measures for null hypothesis significance testing. Many measures have been  
65 developed to describe the structure (Poisot *et al.* (2016)) and diversity (Ohlmann *et al.* (2019), Godsoe *et al.*  
66 (2022)) of probabilistic webs. These models and measures support the use of this approach for the study of a  
67 wide range of ecological questions, from making better predictions of species distribution (Cazelles *et al.*  
68 (2016)) to forecasting the impact of climate change on ecological webs (Gilman *et al.* (2010)).

69 The lack of clear guidelines on the use of probabilistic interaction data is worrisome (Brimacombe *et al.*  
70 (2023)), both for data producers and re-users who generate and manipulate these numbers. Besides  
71 methodological difficulties that may arise when assessing probabilistic interactions, a precise definition of  
72 probabilistic interactions appears to be lacking, making the estimation and use of these data more difficult. We  
73 aim to take a step back by outlining different ways in which probabilistic interactions are defined and used in  
74 network ecology. We distinguish two broad categories of probabilistic webs that necessitate distinct approaches  
75 when applied to key ecological questions: local webs of realized interactions, and regional webs (metawebs) of  
76 potential interactions. We highlight the distinctions in the ecological meaning of these two representations and  
77 show that they yield different statistical outcomes regarding e.g. the spatial and temporal scaling of interactions  
78 and the prediction of binary webs across space. Moreover, there is currently no metadata standard that could  
79 guide the documentation of all types of probabilistic interactions (although see e.g., Salim *et al.* (2022) who  
80 discuss data standards for deterministic mutualistic webs). Well-defined metadata for probabilistic webs would  
81 support more adequate manipulation and integration of interaction data from different sources and guard against  
82 possible misinterpretations arising from ambiguous definitions of probabilistic networks. These metadata  
83 should outline the nature (i.e., local or regional) and type (e.g., predatory or pollination) of the interactions,  
84 provide information regarding the taxonomic level, identities, and characteristics (e.g., life stages) of the  
85 individuals involved in an interaction, present the mathematical formulation of probabilities, including clearly  
86 identified conditional variables (e.g., spatial and temporal scales), and describe the methods and contexts (e.g.,  
87 location, time, environmental conditions) in which interactions were estimated. Inadequately documented  
88 probabilistic interaction data should be used with caution when analyzing ecological webs. Although our  
89 contribution focuses on food webs, our observations and advice can be applied to other types of ecological

90 networks, from plant-pollinator to host-virus networks. Indeed, excluding networks of indirect interactions such  
91 as competition and facilitation networks (Kéfi *et al.* (2015), Kéfi *et al.* (2016)), most ecological webs describe  
92 probabilities of direct interactions, which are conceptually and mathematically analogous regardless of their  
93 biological type (e.g., trophic and mutualistic interactions).

## 94 **Probabilistic representations of interactions**

95 One of the first aspects to take into consideration when estimating or interpreting probabilities of interactions is  
96 knowing if they describe potential or realized interactions. A potential (regional) interaction is defined as the  
97 biological capacity of two taxa to interact (i.e., the probability that they *can* theoretically interact) whereas a  
98 realized (local) interaction is the materialization or observation of this interaction in a well-defined space and  
99 time (i.e., the probability that they interact locally). Here, we use the terms *metaweb* (Dunne (2006)) to  
100 designate regional webs of potential interactions and *local webs* (Poisot *et al.* (2012)) for those of realized  
101 interactions. Metawebs are the network analogs of the species pool, where local webs originate from a subset of  
102 both species (nodes) and interactions (edges) of the regional metaweb (Saravia *et al.* (2022)). Without clear  
103 documentation, it can be challenging to know if published probabilistic webs describe local or regional  
104 interactions, or if so-called probabilities are in reality *interaction scores* (i.e., non-probabilistic quantitative  
105 interactions). When probabilistic regional interactions are used and interpreted as local interactions (and  
106 conversely), this may generate misleading findings during data analysis. We believe that a better understanding  
107 of the differences, similarities, and relationships between these two probabilistic representations of ecological  
108 webs would alleviate interpretation errors and facilitate a more adequate utilization of interaction data.

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

110 Local and metawebs, like any type of network, are made of nodes and edges that can be represented at different  
111 levels of organization and precision. The basic unit of food webs and other ecological networks are individuals  
112 that interact with each other (e.g., by predation, Elton (2001)), forming individual-based networks (Melián *et al.*  
113 (2011)). The aggregation of these individuals into more or less homogeneous groups (e.g., populations, species,  
114 trophic species, families) allows us to represent nodes at broader taxonomic scales, which affects our  
115 interpretation of the properties of these systems (Guimarães (2020)). Moreover, edges linking these nodes can  
116 describe a variety of interaction measures. Ecologists have traditionally represented interactions as binary

117 objects that were considered realized after observing at least one individual from group  $i$  interact with at least  
118 another individual from group  $j$ . Boolean interactions can be viewed as the result of a Bernoulli process  
119  $A_{i,j} \sim \text{Bernoulli}(P(i \rightarrow j))$ , with  $P(i \rightarrow j)$  being the probability of interaction between  $i$  and  $j$  that characterizes  
120 our limited knowledge of the system and its intrinsic spatiotemporal variability. Depending on the type of  
121 network (local or metaweb), the mathematical formulation and interpretation of stochastic parameters like  
122  $P(i \rightarrow j)$  can be linked to environmental and biological factors such as species relative abundance, traits, area,  
123 and time (tbl. 1), for example using logistic regression with a logit link function with continuous explanatory  
124 variables. In these probabilistic network representations in which  $P(i \rightarrow j)$  are edge values, the only two  
125 possible outcomes are the presence ( $A_{i,j} = 1$ ) or absence ( $A_{i,j} = 0$ ) of an interaction between each pair of nodes.  
126 When considering uncertainties around the estimation of  $P(i \rightarrow j)$ , a Beta distribution  $\text{Beta}(\alpha, \beta)$  can also be  
127 used to encompass all possible probability values. Observing an interaction between two taxa at a given  
128 location and time provides important information that can be used to update previous estimates of  $P(i \rightarrow j)$ ,  
129 informing us on the biological capacity of both taxa to interact and the environmental conditions that enabled  
130 them to interact locally.

131 Even though binary webs constitute a highly valuable source of ecological information (Pascual *et al.* (2006)),  
132 they overlook important factors regarding interaction strengths. Represented as quantitative interactions not  
133 confined to the  $[0, 1]$  range, interaction strengths better describe the energy flows, demographic impacts or  
134 frequencies of interactions between nodes (Berlow *et al.* (2004), Borrett & Scharler (2019)), with  $A_{i,j}$  being a  
135 natural  $\mathbb{N}$  or real  $\mathbb{R}$  number depending on the measure. For example, they may represent local interaction rates  
136 between pairs of taxa (e.g., the flower-visiting rates of pollinators in a mutualistic network, Herrera (1989)).  
137 When interaction strengths characterize predation pressure on prey taxa in food webs, they can serve as good  
138 estimators of the parameters describing species interactions in a Lotka-Volterra model (e.g., Emmerson &  
139 Raffaelli (2004)). The extra amount of ecological information in quantitative networks typically comes at a cost  
140 of greater sampling effort and data volume in predictive models (Strydom *et al.* (2021)), which can lead to  
141 relatively high levels of uncertainties when inferring quantitative webs with limited data.  
142 Just like binary networks, the uncertainty and spatiotemporal variability of interaction strengths can be  
143 represented probabilistically. However, the need to estimate the probability distribution of all possible values of  
144 interaction strengths can make the inference of probabilities more challenging in quantitative webs compared to  
145 binary webs, which require only one probability estimate for each interaction. Interaction strengths can follow  
146 various probability distributions depending on the measure used. For instance, they can follow a Poisson

147 distribution  $A_{i,j} \sim \text{Poisson}(\lambda_{i \rightarrow j})$  when predicting frequencies of interactions between pairs of nodes, with  $\lambda_{i \rightarrow j}$   
148 being the expected rate at which individuals of taxa  $i$  and  $j$  interact (e.g., the average number of prey  $j$  consumed  
149 by all predators  $i$  in a given time period). The Poisson distribution can also be 0-inflated after initially modeling  
150 non-interacting taxa (e.g., Boulangeat *et al.* (2012) employ a 0-inflated model to analyze species abundance  
151 following the modeling of species presence and absence), which constitute the majority of taxa pairs in most  
152 local webs due to their typically high sparseness (Jordano (2016)). Because of the methodological difficulties  
153 typically encountered when building deterministic quantitative webs, binary webs, which are easier to sample  
154 (Jordano (2016)) and predict (Strydom *et al.* (2021)), are documented and modeled more frequently in the  
155 literature. Moreover, most published probabilistic networks and methods describe Bernoulli interactions  
156 (whether interaction probabilities are regarded as constant or variable, e.g. represented by Beta distributions),  
157 which underlines the need for better guidelines regarding the interpretation and manipulation of these types of  
158 webs. For these reasons, our primary focus in this contribution will be on addressing the challenges in  
159 interpreting and using Bernoulli interactions, in both probabilistic local and metawebs.

## 160 Local webs: communities interacting in space and time

161 Probabilistic local webs describe how likely taxa are to interact at a given location and time period (i.e.,  
162 interactions are contingent upon the environmental and biological conditions of the community). In local webs,  
163 edges commonly represent our degree of belief that two taxa interact in nature, but can also represent the  
164 probability of *observing* this interaction (Catchen *et al.* (2023)). For example, Kopelke *et al.* (2017) assembled  
165 a dataset of binary local European food webs of willow-galling sawflies and their natural enemies, clearly  
166 referencing each food web in space and time. Gravel *et al.* (2019) used this dataset to infer the probabilities of  
167 locally observing interactions between co-occurring species, with the added aspect of situating local webs  
168 within the context of environmental variables. This was achieved by including temperature and precipitation as  
169 conditional variables in some of their models.

170 We define space as the collection of geographic coordinates  $(x, y, z)$ , with  $(x, y)$  representing longitude and  
171 latitude coordinates, and  $z$  denoting either altitudes or depths. These point coordinates delineate the spatial  
172 boundaries of the system, which may be portrayed as a polyhedron. Ecological interactions may exhibit  
173 variations along altitudinal gradients, as evidenced by changes in hummingbird-plant interactions (Weinstein &  
174 Graham (2017b) and Weinstein & Graham (2017a)) and mosquito biting rates (e.g., Kulkarni *et al.* (2006)) at  
175 different elevations. In contrast, time is defined as the specific time period within which interactions were either

176 observed or predicted. With these definitions in mind, space and time can be conceptualized as distinct patches  
177 or time segments. Treating them as discrete dimensions aligns with the common sampling methods of  
178 ecological webs and provides actual probabilities of interactions across space and time. This differs from the  
179 approach of treating them as continuous variables, where edge values represent probability densities (i.e.,  
180 relative likelihoods of interactions occurring at infinitesimal locations and instants in time) rather than discrete  
181 probability values (which can be obtained by integrating probability densities). By employing discrete locations  
182 and time periods, we can quantify both a volume or area  $A$  and a duration  $t$ , which can be readily used in  
183 spatiotemporal analyses of ecological networks. For example, when studying network-area relationships (NAR,  
184 Galiana *et al.* (2018)), we anticipate that local probabilities of interactions scale positively with area and  
185 duration because taxa have more opportunities to interact as these dimensions expand.

186 The probability that two taxa  $i$  and  $j$  interact locally can be conditional on many environmental and biological  
187 factors. One of these is their co-occurrence  $C_{i,j}$ , which is usually a Boolean describing if the geographic  
188 distributions of both taxa overlap within the study area. Co-occurrence can be modeled probabilistically, in  
189 which case it may conform to a Bernoulli distribution  $C_{i,j} \sim \text{Bernoulli}(P_{i,j}(x,y,z))$ , where the probability of  
190 co-occurrence  $P_{i,j}(x,y,z)$  can be estimated through the application of joint species distribution models (e.g.,  
191 Pollock *et al.* (2014)). Considering that the probability that two non-co-occurring taxa interact locally is zero  
192 (i.e.,  $P_N(i \rightarrow j|C = 0) = 0$ ), the probability of local interaction can be obtained by multiplying the probability of  
193 interaction given co-occurrence with the probability of co-occurrence:

$$P_N(i \rightarrow j) = P_N(i \rightarrow j|C = 1) \times P_{i,j}(x,y,z). \quad (1)$$

194 Other important factors that can impact interaction probabilities at the local scale are taxa relative abundance  
195 (Canard *et al.* (2012)) and traits (Poisot *et al.* (2015)), as well as environmental factors such as temperature  
196 (Angilletta *et al.* (2004)), precipitation (Woodward *et al.* (2012)), habitat structure (Klecka & Boukal (2014)),  
197 and presence of other interacting taxa in the network (Pilosof *et al.* (2017), Kéfi *et al.* (2012)). Here, we use the  
198 variable  $\Omega$  to describe the biological and ecological context in which interaction probabilities were estimated.  
199 For example, if a research team conducts a mesocosm experiment to estimate interaction probabilities between  
200 predators and prey with and without shelters,  $\Omega$  would represent the presence or absence of these shelters. Like  
201 co-occurrence,  $\Omega$  can also be modeled probabilistically when the stochasticity or uncertainty of environmental  
202 and biological factors is considered. In sum,  $\Omega$  represents all ecological and biological variables that were taken

203 into consideration when measuring interaction probabilities and is, therefore, a subset of all factors actually  
 204 impacting ecological interactions.

205 The probability that two taxa  $i$  and  $j$  interact in a local web  $N$  can thus be conditional on the area  $A$ , the time  
 206 interval  $t$ , their co-occurrence  $C$  and chosen environmental and biological conditions  $\Omega$ . This gives us the  
 207 following equation when all of these conditions are included in the estimation of interaction probabilities:

$$P_N(i \rightarrow j|A, t, C, \Omega). \quad (2)$$

208 The representation of the local context in which probabilities are estimated and the variables that should be  
 209 taken into consideration depend on the study system, the objectives of the study, and the resources available to  
 210 the researchers. In other words, these variables do not systematically need to be accounted for. For example, in  
 211 Gravel *et al.* (2019), for the purpose of model comparison, some models incorporated environmental variables  
 212 as conditional factors to estimate interaction probabilities, while others did not. When accounted for, these  
 213 variables should be clearly described in the documentation of the data (Brimacombe *et al.* (2023)),  
 214 preferentially in mathematical terms to avoid any confusion in their interpretation and to limit manipulation  
 215 errors during their re-use. For instance, ecologists should be explicit about their consideration of co-occurrence  
 216 in their estimation of local interaction probabilities. Indeed, it is important to specify if probability values are  
 217 conditional  $P_N(i \rightarrow j|C = 1)$  or not  $P_N(i \rightarrow j)$  on co-occurrence since this can significantly impact the  
 218 interpretation and analysis of the data. In [tbl. 1](#), we present examples of studies that used these diverse  
 219 formulations of probabilistic interactions and conditional variables.

**Table 1: Notation of probabilistic interactions.** The probability of interaction between two taxa  $i$  and  $j$  is interpreted differently in a metaweb  $M$  of potential interactions and a local web  $N$  of realized interactions. Each notation includes a different conditional variable, when applicable. An example of a study employing each of these notations and conditional variables is provided, with the specific variables used indicated in parentheses. The study marked with an asterisk has been carried out on binary webs. Note that interaction probabilities can be contingent upon multiple conditional variables, or none at all. Additionally, local interaction probabilities might represent the probability of observing an interaction, not necessarily its actual occurrence.

---

Notation	Type	Description	Reference
$P_M(i \rightarrow j)$	regional	biological feasibility of the interaction	Strydom <i>et al.</i> (2022)
$P_N(i \rightarrow j)$	local	realization of the interaction	Fortuna & Bascompte (2006) (null model)

---

Notation	Type	Description	Reference
$P_N(i \rightarrow j A)$	local	realization of the interaction in a given area	Galiana <i>et al.</i> (2018) *
$P_N(i \rightarrow j t)$	local	realization of the interaction during a given time period	Weinstein & Graham (2017a)
$P_N(i \rightarrow j C)$	local	realization of the interaction given that the taxa co-occur	Gravel <i>et al.</i> (2019)
$P_N(i \rightarrow j \Omega)$	local	realization of the interaction given environmental conditions	Gravel <i>et al.</i> (2019) (temperature and precipitation)

## 220 Metawebs: regional catalogs of interactions

221 Metawebs are networks of potential interactions that have been designed for broad spatial, temporal, and  
 222 taxonomic scales (e.g., species food webs at the continental scale). They represent the probability that taxa can  
 223 biologically interact regardless of their co-occurrence and local environmental conditions. Regional interactions  
 224 are by definition context-independent, i.e. they are not measured at a specific location and time. In contrast with  
 225 probabilistic local webs, which represent the stochasticity of interactions occurring in nature, probabilistic  
 226 metawebs measure our degree of belief in the capacity of two taxa to interact (i.e., the probability that their  
 227 traits could support an interaction in the right conditions). Consequently, although *neutrally* forbidden  
 228 interactions (i.e., improbable interactions between rare species, Canard *et al.* (2012)) tend to have low  
 229 probability values in local webs, they may exhibit a higher probability in the metaweb. Potential interactions  
 230 describe the probability that there exists at least one combination of phenotypes of taxa  $i$  and  $j$  that can interact  
 231 with each other if they were to encounter. To reduce any biases in our calculation of this probability, it is crucial  
 232 to ensure that the set of traits sampled or considered accurately reflects the overall trait distribution in both taxa.  
 233 This enhances our confidence in concluding that the taxa cannot interact when no positive observation has been  
 234 made and increases our capacity to detect an interaction when it is biologically feasible. In addition to pairwise  
 235 interactions, the extent of sampling effort also contributes to shaping our assessment of metaweb properties, as  
 236 sampling a larger number of local webs allows us to capture more regional interactions (McLeod *et al.*  
 237 (2021a)). A probability of potential interaction in a metaweb  $M$  can be expressed as

$$P_M(i \rightarrow j), \quad (3)$$

which, in contrast with eq. 2, is not conditional on any spatial, temporal, or environmental variables (tbl. 1).

Starting from a selected set of taxa, which are usually distributed within a broad region of interest, probabilistic metawebs can be built using different data sources, including literature review, fieldwork, and predictive models (e.g., the metaweb of trophic interactions between Canadian mammals inferred by Strydom *et al.* (2022)).

Every pair of taxa that have confidently been observed to interact 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 is not the case in local probabilistic webs, in which interaction events usually remain stochastic (i.e.,  $P_N(i \rightarrow j) < 1$ ) after empirically observing interactions because of their intrinsic spatiotemporal variability. Similarly, although rare interactions typically have low probabilities in local webs, they can have high probabilities in metawebs when the traits of both taxa are congruent. On the other hand, interactions that were never observed can have low probability values in both local and metawebs, going as low as 0 for forbidden links. However, because of observation errors due to taxonomic misidentifications and ecological misinterpretations (e.g., due to cryptic species and interactions, Pringle & Hutchinson (2020)), many observations of interactions are in reality false positives. Likewise, forbidden interactions can be false negatives in metawebs, e.g. if they have been assessed for specific phenotypes, locations or time. Employing Bayesian models, whether they are mechanistic or phenomenological, could improve the accuracy of our estimation of interaction probabilities within both local and regional webs (e.g., Bartomeus *et al.* (2016), Cirtwill *et al.* (2019)). This improvement is achieved in metawebs by leveraging prior information regarding the feasibility of interactions along with the empirical data on observed interactions.

## Properties of probabilistic

Probabilistic local and metawebs differ in their type of interactions (i.e., realized or potential) and in the conditional variables upon which interaction values depend. These differences are significant as they influence the characteristics of probabilistic networks. Neglecting to consider them may result in misleading results and interpretation errors when analyzing the properties of probabilistic webs, which could be particularly problematic when addressing crucial ecological questions about networks. Here we compare the characteristics of local and metawebs through the presentation of four common applications of probabilistic interactions. All

263 code and data to reproduce these analyses are available at the Open Science Framework (TBD).

264 **Host-parasite network data**

265 We use the collection of tripartite host-parasite networks sampled across Europe, created by Kopelke *et al.*  
266 (2017), in most of our case studies. This dataset contains well-resolved binary local interactions between  
267 willows (52 species), willow-galling sawflies (96 species), and their parasitoids (126 species). Given its  
268 replicated webs spanning large spatiotemporal scales, this dataset is well-suited for analyzing a variety of  
269 ecological hypotheses and processes. Out of a total of 374 local webs, we retained those containing at least 5  
270 species, resulting in a set of 233 georeferenced local webs. We built a binary metaweb by aggregating all local  
271 interactions, which gave us a regional web composed of 274 species and 1080 interactions. In the first two  
272 panels of fig. 1, we show how the dissimilarity of interactions between common species ( $\beta_{OS}$ ) and the  
273 dissimilarity in species composition ( $\beta_S$ ) between the metaweb and aggregated local webs (Poisot *et al.* (2012))  
274 vary with the number of sampled local webs. This shows that networks of local interactions are highly  
275 dissimilar from the metaweb, both in terms of species and interactions, especially when only a limited number  
276 of sites has been sampled.

277 [Figure 1 about here.]

278 We converted these binary networks into probabilistic ones using models based on simple assumptions. Our  
279 models do not aim to precisely estimate the actual values of probabilistic interactions but rather to generate  
280 simplified networks as illustrative examples to underscore the differences between probabilistic local and  
281 metawebs. We created two probabilistic metawebs by employing constant false positive and false negative rates  
282 for all regional interactions. In the first metaweb, we set both false positive and false negative rates to zero to  
283 prevent artificially inflating the total number of links, enabling a more accurate comparison with binary webs.  
284 This gave us a probability of regional interaction of 1 when at least one interaction has been observed and of 0  
285 in the absence of any observed interaction between a given pair of species. In the second metaweb, we  
286 introduced a 5% false positive rate to account for spurious interactions and a 10% false negative rate to capture  
287 the elevated occurrence of false negatives in ecological networks (Catchen *et al.* (2023)). Observed interactions  
288 were thus given a probability of regional interaction of 95%, whereas unobserved ones were assigned a  
289 probability of 10%.

290 To build probabilistic local webs, we first recognize that local interactions must initially be biologically feasible  
 291 before occurring at a specific time and space. A local probability of interaction  $P_N(i \rightarrow j)$  can be expressed as  
 292 the product of the probability of local interaction given that the two taxa can potentially interact  
 293  $P_N(i \rightarrow j|M_{i \rightarrow j} = 1)$ , which we denote as  $p$  for the sake of simplicity, with their probability of regional  
 294 interaction  $P_M(i \rightarrow j)$ :

$$P_N(i \rightarrow j) = P_N(i \rightarrow j|M_{i \rightarrow j} = 1) \times P_M(i \rightarrow j). \quad (4)$$

295 We built the probabilistic local webs from the binary ones by using the probabilistic metawebs and a constant  
 296 value of  $p$  across interactions. Low values of  $p$  indicate that feasible interactions rarely occur locally,  
 297 intermediate values around 50% suggest considerable spatiotemporal variability, while high values indicate that  
 298 regional interactions are nearly always realized locally. Following eq. 4, the local probability of interaction  
 299 between a given pair of taxa consistently remained equal to or below their probability of regional interaction.

300 In the last two panels of fig. 1, we show how the aggregated number of links and connectance (i.e., the  
 301 proportion of all of the  $S^2$  possible links that are realized, where  $S$  represents species richness) scale with the  
 302 number of sampled local probabilistic webs, according to different values of  $p$ . When aggregating local  
 303 probabilistic webs, the constancy of the probability of regional interaction across the entire study area means  
 304 that any rise in the probability of local interaction is solely attributable to an increase in  $p$ . The probability  $p_{1,2}$ ,  
 305 of local interaction among potentially interacting species in an aggregated web  $N_{1,2}$  is obtained by:

$$p_{1,2} = 1 - (1 - p_1) \times (1 - p_2), \quad (5)$$

306 where  $p_1$  and  $p_2$  are the probabilities of local interaction among two potentially interacting species in the  
 307 subnetworks  $N_1$  and  $N_2$ , respectively.

308 By comparing the scaling relationships observed in binary and probabilistic webs, fig. 1 illustrates that high  
 309 values of  $p$  lead to systematic overestimations in the number of links and connectance, especially when  $p = 1$   
 310 (corresponding to the scenario where local probabilities of interactions are equivalent to the probabilities of  
 311 regional interactions). However, these biases tend to diminish as the number of sampled webs increases.

312 **Taxonomic scaling of interactions**

313 Probabilistic webs offer a versatile approach to tackle a broad array of ecological questions, depending on their  
314 level of organization. For instance, the assemblage of interactions across ecological scales can be explored  
315 through species-based networks, while clade-based networks provide insights into macroevolutionary processes  
316 (e.g., Gomez *et al.* (2010)). Given that our interpretation of the properties and dynamics of ecological webs  
317 depends on their taxonomic scale (Guimarães (2020)), investigating the phylogenetic scaling of network  
318 structure emerges as a promising research avenue. Examining the same system at various taxonomic scales can  
319 yield meaningful and complementary ecological information, and, in our perspective, employing probabilistic  
320 webs is an effective approach for such analyses.

321 There are no inherent differences in the taxonomic scaling between local and metawebs, as only the nodes are  
322 defined taxonomically. In other words, the probability values of edges in both local (eq. 2) and metawebs (eq. 3)  
323 are not conditional on any taxonomic scale. The phylogenetic scale is tied to the definition of the event itself  
324 (i.e., the interaction between two taxa), not to the conditional variables. In both types of webs, transitioning to a  
325 broader level of organization can be accomplished directly by using probabilities from finer scales. For  
326 example, in a network with  $n_A$  species from genus A and  $n_B$  species from genus B, one can compute the  
327 probability that at least one species from genus A interacts with at least one species from genus B as follows:

$$P(A \rightarrow B) = 1 - \prod_{i=1}^{n_A} \prod_{j=1}^{n_B} (1 - P(A_i \rightarrow B_j)), \quad (6)$$

328 where  $A_i$  and  $B_j$  are the species of the corresponding genus. If it is known that at least two of these species  
329 interact (i.e.,  $P(A_i \rightarrow B_j) = 1$  for at least one pair of  $i, j$ ), it implies a probability of genus interaction equal to 1.  
330 Canard *et al.* (2012) built a species-based network following a similar approach, by using simulated interactions  
331 between individuals derived from a neutral model (i.e., a model that assumed ecological equivalence among  
332 individuals). In contrast, a more sophisticated approach is necessary when transitioning from a broader to a  
333 finer level of organization. This is because knowledge of an interaction between two genera does not guarantee  
334 that all possible pairwise combinations of their species will also interact. One possible method is to build a  
335 finer-scale network by generating probabilities of interactions through random sampling from a beta  
336 distribution, parameterized by the broader-scale network.

337 Ideally, our biological interpretation of probabilistic interactions should remain consistent across a network

even if it incorporates heterogeneous levels of organization (e.g., a network whose nodes are composed of both species and trophic species). This is common in ecological webs where taxonomic resolution is typically low (Hemprich-Bennett *et al.* (2021), Vázquez *et al.* (2022)). Interaction probabilities at broader taxonomic scales should be based on probabilities of interactions between individuals, for both local and metawebs. For instance, in local individual-based food webs, the probability that two individuals interact reflects our degree of belief that one individual will consume the other. Similarly, in local species-based food webs, the probability that two species interact represents our degree of belief that at least one individual from the predator species consumes at least another individual from the prey species. Furthermore, in local clade-based food webs, the probability that two clades interact represents our degree of belief that at least two species from these clades interact with each other or, equivalently, that at least two individuals from these clades interact with each other. Fundamentally, the taxonomic scaling of interactions involves aggregating interactions between individuals into larger groups, which may exhibit varying degrees of homogeneity based on the organisms and the study system. In that regard, taxonomic scaling is analogous to the spatial and temporal scaling of interactions, as they all represent different ways to aggregate individuals into broader groups (either spatially, temporally, or taxonomically).

## 352 Spatial and temporal scaling of interactions

353 The investigation of network-area relationships (NARs) and interaction accumulation curves (IACs) is an  
354 important area of research in network ecology. First, NARs elucidate the scaling of network properties (such as  
355 modularity and trophic chain lengths) with spatial scale (e.g., Wood *et al.* (2015), Galiana *et al.* (2018)). The  
356 variations in network structure across spatial scales may stem from the scaling of species richness (species-area  
357 relationships, SARs) and the number of interactions (Brose *et al.* (2004)) with the sampled area. Additionally,  
358 ecological processes unfolding at distinct spatial scales, such as the spatial variability in local community  
359 composition resulting from different sequences of extinction and colonization events, can also contribute to this  
360 variation (Galiana *et al.* (2018)). Next, IACs describe the scaling of the number of observed interactions with  
361 sampling effort (Jordano (2016)), which can also impact various measures of network structure such as  
362 connectance (McLeod *et al.* (2021b)). Apart from sampling effort, the temporal scaling of interactions  
363 elucidates how network structure changes with the duration of the network (Poisot *et al.* (2012)),  
364 acknowledging that distinct interactions take place over time. Given that probabilistic local webs can quantify  
365 the spatiotemporal variability of interactions (eq. 2), they serve as valuable tools for developing and testing  
366 equations describing the scaling of network structure with space and time.

367 Meta and local webs exhibit distinct relationships with spatial and temporal scales. On one hand, probabilistic  
368 metawebs, being independent of any local context, feature regional interactions that do not scale with space and  
369 time. This is because regional interactions depend solely on the biological capacity of two taxa to interact,  
370 regardless of their co-occurrence and specific environmental conditions. This implies that the probability of two  
371 taxa potentially interacting should theoretically be the same in all metawebs in which they are present, provided  
372 that the data and methods used for estimation are consistent. As a result, if a smaller metaweb  $M_1$  is derived  
373 from subsampling a metaweb  $M_0$  to represent potential interactions among a subset of taxa in a smaller region,  
374 their probabilities of interaction should be identical regardless of scale, i.e.  $P_{M_1}(i \rightarrow j) = P_{M_0}(i \rightarrow j)$ . However,  
375 because the number of taxa may be higher in the larger metaweb, it is expected that the overall number of  
376 interactions would be higher in the larger web, despite pairwise probabilities being identical.

377 On the other hand, local interactions scale both spatially and temporally, given that they have more  
378 opportunities to be realized and observed in larger areas and longer durations. This is attributed to factors such  
379 as a higher number of individuals, greater trait variations, and increased opportunities for encounters, as  
380 highlighted by McLeod *et al.* (2020). For example, if a local probabilistic web  $N_1$  with an area  $A_1$  is derived  
381 from a larger web  $N_0$  with an area  $A_0$ , and  $A_1$  is entirely nested within  $A_0$ , interaction probabilities should be  
382 lower in the smaller web, i.e.  $P_{N_1}(i \rightarrow j|A_1 < A_0) \leq P_{N_0}(i \rightarrow j|A_0)$ . However, if  $A_1$  and  $A_0$  are disjoint,  
383 interaction probabilities could be higher in the smaller area, contingent upon their environmental and biological  
384 conditions. Likewise, interaction probabilities are expected to be lower in webs with shorter durations when  
385 time intervals are nested. In fig. 2, we show how the expected number of local host-parasite interactions scales  
386 with area, represented as an expanding latitudinal window, in comparison with regional interactions. Even  
387 though we employed local probabilities of interactions equal to regional interactions for the purpose of  
388 comparison (i.e., using  $p = 1$  here), we notice that the total number of regional interactions scales more rapidly  
389 than local interactions. This is because numerous regional interactions involve species that never co-occur, and  
390 as a result, these interactions are not captured in local webs.

391 [Figure 2 about here.]

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

393 Predicting local webs across time and space is a pivotal goal of network ecology (Strydom *et al.* (2021)). In a  
394 context of scarcity of interaction data (Jordano (2016)), ecologists must resort to predictive models for

395 reconstructing networks at fine spatial and temporal scales. For instance, real-time biomonitoring data coupled  
 396 with appropriate numerical models (Bohan *et al.* (2017)) can be employed to reconstruct local ecological webs,  
 397 opening avenues for in-depth studies on local ecosystem functioning and dynamics. Apart from their predictive  
 398 applications, statistical models can also be crafted for descriptive purposes, elucidating key parameters of  
 399 interest such as probabilities of interactions. In such instances, the parameter values themselves offer valuable  
 400 ecological insights. Various types of models (e.g., Bayesian and machine learning models) have been used for  
 401 both predictive and descriptive purposes in the realm of ecological interactions (Strydom *et al.* (2021)). The  
 402 probabilistic representation of interactions acknowledges the inherent uncertainty in these models, typically  
 403 expressed through probability distributions. We introduce and develop a simple generative mechanistic model  
 404 for probabilistic local interactions that takes into consideration their spatiotemporal variability (i.e. a  
 405 spatiotemporally explicit model of local interactions). It is essential to note that our model is not designed for  
 406 regional interactions, which are scale-independent. Rather, it could prove valuable for predicting local  
 407 interactions across time and space by generating new interaction data following parameter inference.

408 As indicated by eq. 1, the probability that two taxa  $i$  and  $j$  interact at a specific location  $(x, y, z)$  is determined by  
 409 the product of their probability of interaction given co-occurrence and their probability of co-occurrence. First,  
 410 their co-occurrence probability can be calculated using their individual occurrence probabilities  $P_i(x, y, z)$  and  
 411  $P_j(x, y, z)$ , along with the strength of association  $\gamma$  between their occurrences and co-occurrence (Cazelles *et al.*  
 412 (2016)):

$$P_{i,j}(x, y, z) = P_i(x, y, z)P_j(x, y, z)\gamma. \quad (7)$$

413 When  $\gamma > 1$ , it signifies a positive association in the geographic distributions of both taxa, indicating that the  
 414 presence of one taxon enhances the probability of occurrence of the other. In empirical webs,  $\gamma > 1$  holds true  
 415 for the majority of species pairs (Catchen *et al.* (2023)). The co-occurrence of both taxa is modeled as the  
 416 outcome of a Bernoulli trial

$$C \sim Bernoulli(P_{i,j}(x, y, z)). \quad (8)$$

417 Next, the probability of interaction given co-occurrence can be made temporally explicit by modeling it as a  
 418 Poisson process with rate parameter  $\lambda$ . This parameter represents the expected frequency of interaction between

419 the taxa within a defined time interval and can be estimated using prior data on interaction strengths, when  
 420 accessible. The probability that two co-occurring taxa engage in an interaction during a time period  $t_0$  is given  
 421 by:

$$P_N(i \rightarrow j|C = 1) = 1 - e^{-\lambda t_0}, \quad (9)$$

422 which tends toward 1 as  $t_0 \rightarrow \infty$ . It is important to note that the units of  $\lambda$  and  $t_0$  are complementary. For  
 423 instance, if the duration  $t_0$  is measured in months,  $\lambda$  denote the expected number of interactions per month.

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

$$I \sim \text{Bernoulli}(P_N(i \rightarrow j)) \quad (10)$$

427

$$P_N(i \rightarrow j) = P_i(x, y, z)P_j(x, y, z)\gamma(1 - e^{-\lambda t_0}) \quad (11)$$

$$\gamma \sim \text{Gamma}(2, 0.5) \quad (12)$$

$$\lambda \sim \text{Exponential}(2) \quad (13)$$

428 In fig. 3, we show the variation in the probability of interaction under different parameter values. In the right  
 429 panel, we notice that, irrespective of the interaction rate  $\lambda$ , the probability of interaction converges toward an  
 430 asymptote determined by the co-occurrence  $P_{i,j}(x, y, z)$  (eq. 7). This model can be customized in different ways,  
 431 such as by linking  $\lambda$  with specific environmental variables or explicitly incorporating observation errors (i.e.,  
 432 the probabilities of false negatives and false positives).

433 [Figure 3 about here.]

434 **Sampling binary webs**

435 The prediction of binary interactions through random draws is an important application of probabilistic webs.  
436 This approach proves beneficial for analyzing the structural characteristics of probabilistic webs, particularly in  
437 the absence of specific analytical measures. By simulating independent Bernoulli trials for each interaction, a  
438 binary network can be generated. A probability distribution of network properties can be obtained by measuring  
439 network structure across multiple randomly generated networks (Poisot *et al.* (2016)). This method enables the  
440 representation of the variability of network structure, albeit with possible biases when connectance is low  
441 (Poisot & Gravel (2014), Chagnon (2015)). Employing this strategy to generate binary networks under a null  
442 model facilitates null hypothesis significance testing, wherein the observed measure is compared against the  
443 simulated distribution (e.g., Bascompte *et al.* (2003)). Additionally, randomly generating binary networks  
444 across spatial and temporal dimensions aids in representing the spatiotemporal variability of network structure,  
445 allowing the testing of ecological hypotheses regarding interactions on large spatial and temporal scales.

446 There are at least two distinct approaches to sample binary networks from probabilistic webs across space, for  
447 example, when attempting to predict a binary network for each of a number of locations within a given region.  
448 The first approach involves performing a singular Bernoulli trial for each pair of taxa based on their regional  
449 probability of interaction. In employing this approach, every pair of taxa predicted to interact in the binary  
450 metaweb will be treated as interacting in all local webs where they co-occur. This will result in local pairwise  
451 interactions without spatial variation. The second approach is to independently sample each of the local  
452 probabilistic webs, which can be achieved by first generating distinct probabilistic networks for each location.  
453 These local probabilistic webs may vary in taxa composition and interaction probabilities. Subsequently, binary  
454 networks can be independently sampled for each location, introducing spatial variation in binary interactions.

455 In fig. 4, we compare the average connectance of binary networks resulting from these two sampling techniques,  
456 where regional and local interactions are drawn from our host-parasite probabilistic networks, generating a  
457 number of binary web realizations for each site in the dataset. These two sampling approaches yield different  
458 outcomes, particularly for lower values of  $p$ , which denote instances where regional interactions do not  
459 consistently manifest locally. Small discrepancies are also apparent between these techniques when we equate  
460 the probability of local interaction to the probability of regional interaction (i.e., when using  $p = 1.0$  in eq. 4),  
461 especially when the number of binary network samples for each location is low. As anticipated, we observe that  
462 sampling binary interactions from the metaweb tends to overestimate connectance on average compared to

463 sampling them from local webs. Furthermore, we observe an increase in the variability of connectance when  
464 employing a single sample, representing what we consider as a more tangible process leading to the realization  
465 of local and regional interactions in nature.

466 [Figure 4 about here.]

467 The choice of a sampling approach can influence the selection of grid cell size when delineating local  
468 communities within a broader region of interest. In the first approach, pairwise interactions remain constant  
469 irrespective of cell size since they are sampled only once from the metaweb. However, in the second approach,  
470 local interaction probabilities are contingent on the network area. For instance, consider the local webs  $N_1$  and  
471  $N_2$  with an area of  $\frac{1}{2}A_0$ , both nested within  $A_0$  but disjoint from each other, forming  $N_0$ . If we treat  $N_1$  and  $N_2$   
472 as independent, the probability of interaction between taxa  $i$  and  $j$  in  $N_0$  is given by:

$$P_{N_0}(i \rightarrow j) = 1 - (1 - P_{N_1}(i \rightarrow j)) \times (1 - P_{N_2}(i \rightarrow j)). \quad (14)$$

473 Due to its larger area, the probability that the two taxa interact in  $N_0$  is equal or greater than in  $N_1$  and  $N_2$ .  
474 When sampling binary interactions from local webs, it is crucial to sample at the same spatial scale for which  
475 probabilities were estimated. Otherwise, interaction probabilities must be adjusted to align with the intended  
476 cell size, preventing systematic biases in predictions.

## 477 Prediction of local webs from metawebs

478 Metawebs serve as a valuable source of ecological information for predicting local webs across time and space.  
479 Binary local webs can be reconstructed by selecting a subset of taxa and interactions from the metaweb (Dunne  
480 (2006)). This implies that metawebs consistently contain more interactions than their corresponding local webs.  
481 Determining the list of taxa to select can be achieved empirically or through methods like range maps or species  
482 distribution models. As the species composition of a community is arguably less difficult to sample or predict  
483 than its interactions, the primary challenge lies in deciding which interactions to select from the metaweb, a task  
484 that may necessitate advanced statistical models and ecological expertise. Inferring the structure of local webs  
485 from the metaweb before predicting specific local pairwise interactions could hold promise (Strydom *et al.*  
486 (2021)), considering that the structure of local webs is constrained by the metaweb (Saravia *et al.* (2022)).

487 Building probabilistic local webs from a probabilistic metaweb involves a reduction in the value of pairwise  
 488 interaction probabilities. This decrease is attributed to the prerequisite that two taxa must initially possess the  
 489 capacity to interact before engaging in local interactions (eq. 4). Therefore, inferring local webs from their  
 490 metaweb while maintaining identical interaction probability values would introduce systematic biases into the  
 491 predictions. In such cases, these networks would essentially represent smaller-scale metawebs of potential  
 492 interactions, possibly leading to misinterpretations by being perceived as local interactions. As proposed by  
 493 McLeod *et al.* (2021b), although metawebs do not capture the spatiotemporal variability of interactions, they  
 494 establish an upper limit for local interactions. In other words, the probability of two taxa interacting at a specific  
 495 location and time is consistently lower or equal to the probability of their regional interaction:

$$P_N(i \rightarrow j|A, t, C, \Omega) \leq P_M(i \rightarrow j). \quad (15)$$

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

$$\int_{\Omega} \int_A \int_t P_N(i \rightarrow j|A, t, \Omega) dt dA d\Omega \leq P_M(i \rightarrow j). \quad (16)$$

500 Estimating more precisely the probability  $p$  that two taxa interact locally if they can potentially interact allows  
 501 for improved predictions of local webs from a probabilistic metaweb. This task is challenging due to the  
 502 variability of this probability across space and time, as well as its variability across pairwise interactions within  
 503 a network. Using simple models of  $p$ , as demonstrated in our case studies, represents an initial step toward the  
 504 overarching objective of reconstructing local webs from metawebs.

## 505 Conclusion

506 In this contribution, we underline the importance of network metadata for adequately interpreting and  
 507 manipulating probabilistic interaction data. The mathematical representation of probabilities and their statistical  
 508 properties depend on the type of interactions (local or regional) and the conditions under which these  
 509 interactions were evaluated. We showed that probabilistic local and metawebs differ in their relationship to

510 spatial and temporal scales, with regional interactions remaining consistent across scales. In contrast with  
511 metawebs, local interactions are measured in a specific context (e.g., in a given area, time, and biological and  
512 environmental conditions) and depend on taxa co-occurrence. These important conceptual differences bring to  
513 light the need to use probabilistic data with caution, for instance when generating binary network realizations  
514 across space and predicting local webs from metawebs. Clear metadata describing the type of interaction and  
515 the variables used in their estimation are required to ensure adequate data manipulation. Better data practices  
516 and foundations for probabilistic thinking in network ecology could enable more reliable assessments of the  
517 spatiotemporal variability and uncertainty of biotic interactions.

518 It is essential to enhance our comprehension of both regional and local interactions, especially considering the  
519 current scarcity of interaction data. However, while sampling biological communities does decrease the  
520 uncertainty of interactions by accumulating evidence for their feasibility and local realization, there is a limit to  
521 how much we can diminish their inherent randomness. In metawebs, probabilities reflect our limited knowledge  
522 of interactions (i.e., our degree of belief that interactions are feasible), which is expected to improve with a  
523 larger volume of data. We should anticipate that regional interactions will become more definitive (with  
524 probabilities approaching 0 or 1) as we investigate various conditions, including different combinations of  
525 species traits. Conversely, in the case of local webs, which can be seen as random instances of metawebs,  
526 randomness cannot be as reduced. In local interactions, probabilities can be divided into a component  
527 representing uncertainty and another representing spatiotemporal variability. Owing to environmental  
528 heterogeneity, there will invariably be instances in which an interaction occurs and others in which it does not,  
529 across different times and locations, irrespective of the extent to which we can improve our knowledge of its  
530 biological feasibility and the local conditions that facilitate its occurrence. When local webs depict probabilities  
531 of observing interactions rather than just their actual occurrence, we must also consider the observation  
532 variability as an additional source of randomness. Every ecological process is stochastic but there is also a  
533 possibility that a phenomenon goes undetected. Quantifying and partitioning this stochasticity will enable us to  
534 make more accurate predictions about ecological interactions at various spatial and temporal scales. This will  
535 prove to be of vital importance as our time to understand nature runs out, especially at the places where the  
536 impacts of climate change and habitat loss hit harder.

537 **Acknowledgment**

538 We acknowledge that this study was conducted on land within the traditional unceded territory of the Saint  
539 Lawrence Iroquoian, Anishinabewaki, Mohawk, Huron-Wendat, and Omàmiwininiwak nations. This work was  
540 supported by the Institute for Data Valorisation (IVADO) and the Natural Sciences and Engineering Research  
541 Council of Canada (NSERC) Collaborative Research and Training Experience (CREATE) program, through the  
542 Computational Biodiversity Science and Services (BIOS<sup>2</sup>) program. A special thanks to all members of the  
543 Black Holes and Revelations working group (organized by BIOS<sup>2</sup>) for their insightful discussions and valuable  
544 feedback on this manuscript.

545 **References**

- 546 Abrego, N., Roslin, T., Huotari, T., Ji, Y., Schmidt, N.M., Wang, J., *et al.* (2021). [Accounting for species](#)  
547 [interactions is necessary for predicting how arctic arthropod communities respond to climate change.](#)  
548 *Ecography*, 44, 885–896.
- 549 Andrade-Pacheco, R., Rerolle, F., Lemoine, J., Hernandez, L., Meité, A., Juziwelo, L., *et al.* (2020). [Finding](#)  
550 [hotspots: Development of an adaptive spatial sampling approach.](#) *Scientific Reports*, 10, 10939.
- 551 Angilletta, M.J., Jr., Steury, T.D. & Sears, M.W. (2004). [Temperature, Growth Rate, and Body Size in](#)  
552 [Ectotherms: Fitting Pieces of a Life-History Puzzle](#)<sup>1</sup>. *Integrative and Comparative Biology*, 44, 498–509.
- 553 Araujo, M.B., Rozenfeld, A., Rahbek, C. & Marquet, P.A. (2011). [Using species co-occurrence networks to](#)  
554 [assess the impacts of climate change.](#) *Ecography*, 34, 897–908.
- 555 Bartomeus, I., Gravel, D., Tylianakis, J.M., Aizen, M.A., Dickie, I.A. & Bernard-Verdier, M. (2016). [A](#)  
556 [common framework for identifying linkage rules across different types of interactions.](#) *Functional Ecology*,  
557 30, 1894–1903.
- 558 Bascompte, J., Jordano, P., Melián, C.J. & Olesen, J.M. (2003). [The nested assembly of plantanimal mutualistic](#)  
559 [networks.](#) *Proceedings of the National Academy of Sciences*, 100, 9383–9387.
- 560 Berlow, E.L., Neutel, A.-M., Cohen, J.E., De Ruiter, P.C., Ebenman, B., Emmerson, M., *et al.* (2004).  
561 [Interaction strengths in food webs: Issues and opportunities.](#) *Journal of Animal Ecology*, 73, 585–598.

- 562 Blanchet, F.G., Cazelles, K. & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions.  
563 *Ecology Letters*, 23, 1050–1063.
- 564 Bohan, D.A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A.J. & Woodward, G. (2017).  
565 Next-Generation Global Biomonitoring: Large-scale, Automated Reconstruction of Ecological Networks.  
566 *Trends in Ecology & Evolution*, 32, 477–487.
- 567 Bolnick, D.I., Amarasekare, P., Araújo, M.S., Bürger, R., Levine, J.M., Novak, M., *et al.* (2011). Why  
568 intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution*, 26, 183–192.
- 569 Borrett, S.R. & Scharler, U.M. (2019). Walk partitions of flow in Ecological Network Analysis: Review and  
570 synthesis of methods and indicators. *Ecological Indicators*, 106, 105451.
- 571 Boulangeat, I., Gravel, D. & Thuiller, W. (2012). Accounting for dispersal and biotic interactions to disentangle  
572 the drivers of species distributions and their abundances. *Ecology Letters*, 15, 584–593.
- 573 Brimacombe, C., Bodner, K., Michalska-Smith, M., Poisot, T. & Fortin, M.-J. (2023). Shortcomings of reusing  
574 species interaction networks created by different sets of researchers. *PLOS Biology*, 21, e3002068.
- 575 Broom, M. & Ruxton, G.D. (2005). You can run or you can hide: Optimal strategies for cryptic prey against  
576 pursuit predators. *Behavioral Ecology*, 16, 534–540.
- 577 Brose, U., Ostling, A., Harrison, K. & Martinez, N.D. (2004). Unified spatial scaling of species and their  
578 trophic interactions. *Nature*, 428, 167–171.
- 579 Canard, E., Mouquet, N., Marescot, L., Gaston, K.J., Gravel, D. & Mouillot, D. (2012). Emergence of  
580 Structural Patterns in Neutral Trophic Networks. *PLOS ONE*, 7, e38295.
- 581 Catchen, M.D., Poisot, T., Pollock, L.J. & Gonzalez, A. (2023). The missing link: Discerning true from false  
582 negatives when sampling species interaction networks.
- 583 Cazelles, K., Araujo, M.B., Mouquet, N. & Gravel, D. (2016). A theory for species co-occurrence in interaction  
584 networks. *Theoretical Ecology*, 9, 39–48.
- 585 Chagnon, P.-L. (2015). Characterizing topology of ecological networks along gradients: The limits of metrics'  
586 standardization. *Ecological Complexity*, 22, 36–39.
- 587 Choh, Y., Ignacio, M., Sabelis, M.W. & Janssen, A. (2012). Predator-prey role reversals, juvenile experience  
588 and adult antipredator behaviour. *Scientific Reports*, 2, 728.

- 589 Cirtwill, A.R., Eklöf, A., Roslin, T., Wootton, K. & Gravel, D. (2019). [A quantitative framework for](#)  
590 [investigating the reliability of empirical network construction.](#) *Methods in Ecology and Evolution*, 10,  
591 902–911.
- 592 Dallas, T., Park, A.W. & Drake, J.M. (2017). [Predicting cryptic links in host-parasite networks.](#) *PLOS*  
593 [Computational Biology, 13, e1005557.](#)
- 594 Dunne, J.A. (2006). The Network Structure of Food Webs. In: *Ecological networks: Linking structure and*  
595 [dynamics](#) (eds. Dunne, J.A. & Pascual, M.). Oxford University Press, pp. 27–86.
- 596 Elmasri, M., Farrell, M.J., Davies, T.J. & Stephens, D.A. (2020). [A hierarchical Bayesian model for predicting](#)  
597 [ecological interactions using scaled evolutionary relationships.](#) *The Annals of Applied Statistics*, 14,  
598 221–240.
- 599 Elton, C.S. (2001). *Animal Ecology*. University of Chicago Press, Chicago, IL.
- 600 Emmerson, M.C. & Raffaelli, D. (2004). [Predator-prey body size, interaction strength and the stability of a real](#)  
601 [food web.](#) *Journal of Animal Ecology*, 73, 399–409.
- 602 Fortuna, M.A. & Bascompte, J. (2006). [Habitat loss and the structure of plant-animal mutualistic networks.](#)  
603 *Ecology Letters*, 9, 281–286.
- 604 Fu, X., Seo, E., Clarke, J. & Hutchinson, R.A. (2021). [Link Prediction Under Imperfect Detection:](#)  
605 [Collaborative Filtering for Ecological Networks.](#) *IEEE Transactions on Knowledge and Data Engineering*,  
606 33, 3117–3128.
- 607 Galiana, N., Lurgi, M., Claramunt-López, B., Fortin, M.-J., Leroux, S., Cazelles, K., et al. (2018). [The spatial](#)  
608 [scaling of species interaction networks.](#) *Nature Ecology & Evolution*, 2, 782–790.
- 609 Gilman, S.E., Urban, M.C., Tewksbury, J., Gilchrist, G.W. & Holt, R.D. (2010). [A framework for community](#)  
610 [interactions under climate change.](#) *Trends in Ecology & Evolution*, 25, 325–331.
- 611 Godsoe, W., Murray, R. & Iritani, R. (2022). [Species interactions and diversity: A unified framework using Hill](#)  
612 [numbers.](#) *Oikos*, n/a, e09282.
- 613 Golubski, A.J. & Abrams, P.A. (2011). [Modifying modifiers: What happens when interspecific interactions](#)  
614 [interact?](#) *Journal of Animal Ecology*, 80, 1097–1108.
- 615 Gomez, J.M., Verdu, M. & Perfectti, F. (2010). [Ecological interactions are evolutionarily conserved across the](#)  
616 [entire tree of life.](#) *Nature*, 465, 918–U6.

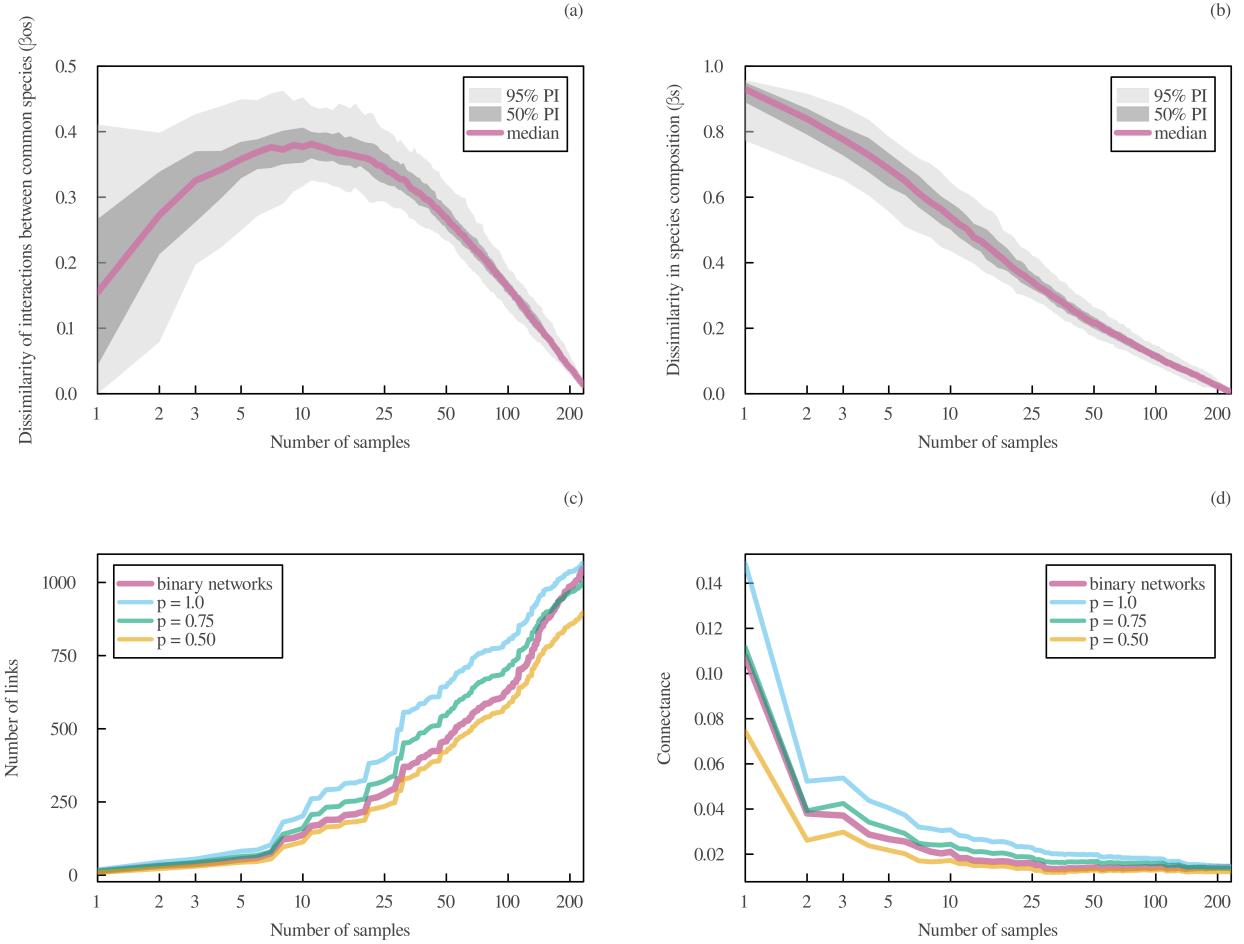
- 617 Gonzalez, A. & Londoño, M.C. (2022). Monitor biodiversity for action. *Science*, 378, 1147–1147.
- 618 Gonzalez-Varo, J.P. & Traveset, A. (2016). The Labile Limits of Forbidden Interactions. *Trends in Ecology &*  
619 *Evolution*, 31, 700–710.
- 620 Gravel, D., Baiser, B., Dunne, J.A., Kopelke, J.-P., Martinez, N.D., Nyman, T., *et al.* (2019). Bringing Elton  
621 and Grinnell together: A quantitative framework to represent the biogeography of ecological interaction  
622 networks. *Ecography*, 42, 401–415.
- 623 Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. (2013). Inferring food web structure from  
624 predator-prey body size relationships. *Methods in Ecology and Evolution*, 4, 1083–1090.
- 625 Guimarães, P.R. (2020). The Structure of Ecological Networks Across Levels of Organization. *Annual Review*  
626 *of Ecology, Evolution, and Systematics*, 51, 433–460.
- 627 Guimerà, R. & Sales-Pardo, M. (2009). Missing and spurious interactions and the reconstruction of complex  
628 networks. *Proceedings of the National Academy of Sciences*, 106, 22073–22078.
- 629 Hemprich-Bennett, D.R., Oliveira, H.F.M., Le Comber, S.C., Rossiter, S.J. & Clare, E.L. (2021). Assessing the  
630 impact of taxon resolution on network structure. *Ecology*, 102, e03256.
- 631 Herrera, C.M. (1989). Pollinator abundance, morphology, and flower visitation rate: Analysis of the “quantity”  
632 component in a plant-pollinator system. *Oecologia*, 80, 241–248.
- 633 Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M. & Ladle, R.J. (2015). Seven Shortfalls  
634 that Beset Large-Scale Knowledge of Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*,  
635 46, 523–549.
- 636 Jordano, P. (1987). Patterns of Mutualistic Interactions in Pollination and Seed Dispersal: Connectance,  
637 Dependence Asymmetries, and Coevolution. *The American Naturalist*, 129, 657–677.
- 638 Jordano, P. (2016). Sampling networks of ecological interactions. *Functional Ecology*, 30, 1883–1893.
- 639 Jordano, P., Bascompte, J. & Olesen, J.M. (2003). Invariant properties in coevolutionary networks of  
640 plant-animal interactions. *Ecology Letters*, 6, 69–81.
- 641 Kéfi, S., Berlow, E.L., Wieters, E.A., Joppa, L.N., Wood, S.A., Brose, U., *et al.* (2015). Network structure  
642 beyond food webs: Mapping non-trophic and trophic interactions on Chilean rocky shores. *Ecology*, 96,  
643 291–303.

- 644 Kéfi, S., Berlow, E.L., Wieters, E.A., Navarrete, S.A., Petchey, O.L., Wood, S.A., *et al.* (2012). [More than a meal... integrating non-feeding interactions into food webs](#). *Ecology Letters*, 15, 291–300.
- 645
- 646 Kéfi, S., Miele, V., Wieters, E.A., Navarrete, S.A. & Berlow, E.L. (2016). [How Structured Is the Entangled Bank? The Surprisingly Simple Organization of Multiplex Ecological Networks Leads to Increased Persistence and Resilience](#). *PLOS Biology*, 14, e1002527.
- 647
- 648
- 649 Klecka, J. & Boukal, D.S. (2014). [The effect of habitat structure on prey mortality depends on predator and prey microhabitat use](#). *Oecologia*, 176, 183–191.
- 650
- 651 Kopelke, J.-P., Nyman, T., Cazelles, K., Gravel, D., Vissault, S. & Roslin, T. (2017). [Food-web structure of willow-galling sawflies and their natural enemies across Europe](#). *Ecology*, 98, 1730–1730.
- 652
- 653 Kulkarni, M.A., Kweka, E., Nyale, E., Lyatuu, E., Mosha, F.W., Chandramohan, D., *et al.* (2006). [Entomological Evaluation of Malaria Vectors at Different Altitudes in Hai District, Northeastern Tanzania](#). *Journal of Medical Entomology*, 43, 580–588.
- 654
- 655
- 656 McCann, K. (2007). Protecting biostructure. *Nature*, 446, 29–29.
- 657
- 658 McCann, K.S. (2011). Food webs (MPB-50). In: *Food Webs (MPB-50)*. Princeton University Press.
- 659
- 660 McLeod, A., Leroux, S.J., Gravel, D., Chu, C., Cirtwill, A.R., Fortin, M.-J., *et al.* (2021a). [Sampling and asymptotic network properties of spatial multi-trophic networks](#). *Oikos*, 130, 2250–2259.
- 661
- 662 McLeod, A., Leroux, S.J., Gravel, D., Chu, C., Cirtwill, A.R., Fortin, M.-J., *et al.* (2021b). [Sampling and asymptotic network properties of spatial multi-trophic networks](#). *Oikos*, n/a.
- 663
- 664 McLeod, A.M., Leroux, S.J. & Chu, C. (2020). [Effects of species traits, motif profiles, and environment on spatial variation in multi-trophic antagonistic networks](#). *Ecosphere*, 11, e03018.
- 665
- 666 Melián, C.J., Vilas, C., Baldó, F., González-Ortegón, E., Drake, P. & Williams, R.J. (2011). [Eco-evolutionary Dynamics of Individual-Based Food Webs](#). In: *Advances in Ecological Research*, The Role of Body Size in Multispecies Systems (ed. Belgrano, A.). Academic Press, pp. 225–268.
- 667
- 668 Ohlmann, M., Miele, V., Dray, S., Chalmandrier, L., O'Connor, L. & Thuiller, W. (2019). [Diversity indices for ecological networks: A unifying framework using Hill numbers](#). *Ecology Letters*, 22, 737–747.
- 669
- 670 Olesen, J.M., Bascompte, J., Dupont, Y.L., Elberling, H., Rasmussen, C. & Jordano, P. (2010). [Missing and forbidden links in mutualistic networks](#). *Proceedings of the Royal Society B: Biological Sciences*, 278, 725–732.
- 671

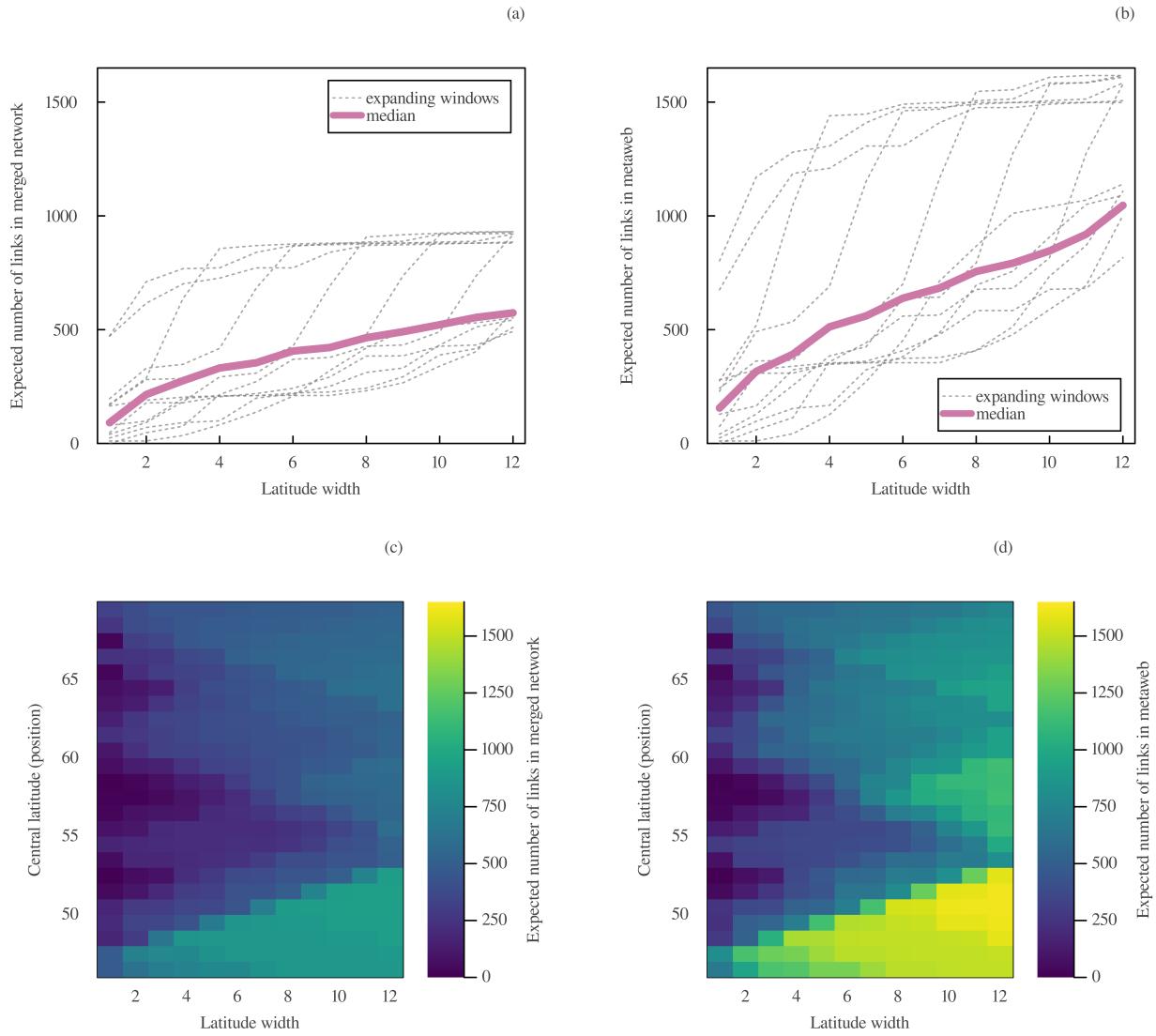
- 672 Pascual, M., Dunne, J.A. & Dunne, J.A. (2006). *Ecological Networks: Linking Structure to Dynamics in Food*  
673 *Webs*. Oxford University Press, USA.
- 674 Petchey, O.L., Beckerman, A.P., Riede, J.O. & Warren, P.H. (2008). [Size, foraging, and food web structure](#).  
675 *Proceedings of the National Academy of Sciences*, 105, 4191–4196.
- 676 Pilosof, S., Porter, M.A., Pascual, M. & Kéfi, S. (2017). [The multilayer nature of ecological networks](#). *Nature*  
677 *Ecology & Evolution*, 1, 1–9.
- 678 Poisot, T., Canard, E., Mouillot, D., Mouquet, N. & Gravel, D. (2012). [The dissimilarity of species interaction](#)  
679 [networks](#). *Ecology Letters*, 15, 1353–1361.
- 680 Poisot, T., Cirtwill, A.R., Cazelles, K., Gravel, D., Fortin, M.-J. & Stouffer, D.B. (2016). [The structure of](#)  
681 [probabilistic networks](#). *Methods in Ecology and Evolution*, 7, 303–312.
- 682 Poisot, T. & Gravel, D. (2014). [When is an ecological network complex? Connectance drives degree](#)  
683 [distribution and emerging network properties](#). *PeerJ*, 2, e251.
- 684 Poisot, T., Stouffer, D.B. & Gravel, D. (2015). [Beyond species: Why ecological interaction networks vary](#)  
685 [through space and time](#). *Oikos*, 124, 243–251.
- 686 Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., *et al.* (2014). [Understanding](#)  
687 [co-occurrence by modelling species simultaneously with a Joint Species Distribution Model \(JSDM\)](#).  
688 *Methods in Ecology and Evolution*, 5, 397–406.
- 689 Pomeranz, J.P.F., Thompson, R.M., Poisot, T. & Harding, J.S. (2019). [Inferring predator-prey interactions in](#)  
690 [food webs](#). *Methods in Ecology and Evolution*, 10, 356–367.
- 691 Pringle, R.M. & Hutchinson, M.C. (2020). [Resolving Food-Web Structure](#). *Annual Review of Ecology,*  
692 *Evolution, and Systematics*, 51, 55–80.
- 693 Proulx, S.R., Promislow, D.E.L. & Phillips, P.C. (2005). [Network thinking in ecology and evolution](#). *Trends in*  
694 *Ecology & Evolution*, SPECIAL ISSUE: BUMPER BOOK REVIEW, 20, 345–353.
- 695 Pulliam, H.R. (1974). [On the Theory of Optimal Diets](#). *The American Naturalist*, 108, 59–74.
- 696 Rooney, N. & McCann, K.S. (2012). [Integrating food web diversity, structure and stability](#). *Trends in Ecology*  
697 *& Evolution*, 27, 40–46.

- 698 Salim, J.A., Saraiva, A.M., Zermoglio, P.F., Agostini, K., Wolowski, M., Drucker, D.P., *et al.* (2022). Data  
699 standardization of plantpollinator interactions. *GigaScience*, 11, giac043.
- 700 Sanders, D. & van Veen, F.J.F. (2012). Indirect commensalism promotes persistence of secondary consumer  
701 species. *Biology Letters*, 8, 960–963.
- 702 Saravia, L.A., Marina, T.I., Kristensen, N.P., De Troch, M. & Momo, F.R. (2022). Ecological network  
703 assembly: How the regional metaweb influences local food webs. *Journal of Animal Ecology*, 91, 630–642.
- 704 Schmeller, D.S., Julliard, R., Bellingham, P.J., Böhm, M., Brummitt, N., Chiarucci, A., *et al.* (2015). Towards a  
705 global terrestrial species monitoring program. *Journal for Nature Conservation*, 25, 51–57.
- 706 Singer, M.C. & McBride, C.S. (2012). Geographic mosaics of species' association: A definition and an  
707 example driven by plantinsect phenological synchrony. *Ecology*, 93, 2658–2673.
- 708 Stouffer, D.B., Rezende, E.L. & Amaral, L.A.N. (2011). The role of body mass in diet contiguity and food-web  
709 structure. *Journal of Animal Ecology*, 80, 632–639.
- 710 Strydom, T., Bouskila, S., Banville, F., Barros, C., Caron, D., Farrell, M.J., *et al.* (2022). Food web  
711 reconstruction through phylogenetic transfer of low-rank network representation. *Methods in Ecology and  
712 Evolution*, 13.
- 713 Strydom, T., Catchen, M.D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., *et al.* (2021). A  
714 roadmap towards predicting species interaction networks (across space and time). *Philosophical  
715 Transactions of the Royal Society B-Biological Sciences*, 376, 20210063.
- 716 Tylianakis, J.M., Tscharntke, T. & Lewis, O.T. (2007). Habitat modification alters the structure of tropical  
717 hostparasitoid food webs. *Nature*, 445, 202–205.
- 718 Vázquez, D.P., Peralta, G., Cagnolo, L., Santos, M. & Igual, E. autores contribuyeron por. (2022). Ecological  
719 interaction networks. What we know, what we don't, and why it matters. *Ecología Austral*, 32, 670–697.
- 720 Weinstein, B.G. & Graham, C.H. (2017a). On comparing traits and abundance for predicting species  
721 interactions with imperfect detection. *Food Webs*, 11, 17–25.
- 722 Weinstein, B.G. & Graham, C.H. (2017b). Persistent bill and corolla matching despite shifting temporal  
723 resources in tropical hummingbird-plant interactions. *Ecology Letters*, 20, 326–335.
- 724 Wood, S.A., Russell, R., Hanson, D., Williams, R.J. & Dunne, J.A. (2015). Effects of spatial scale of sampling  
725 on food web structure. *Ecology and Evolution*, 5, 3769–3782.

- 726 Woodward, G., Benstead, J.P., Beveridge, O.S., Blanchard, J., Brey, T., Brown, L.E., *et al.* (2010). [Chapter 2 -](#)  
727 [Ecological Networks in a Changing Climate](#). In: *Advances in Ecological Research*, Ecological Networks  
728 (ed. Woodward, G.). Academic Press, pp. 71–138.
- 729 Woodward, G., Brown, L.E., Edwards, F.K., Hudson, L.N., Milner, A.M., Reuman, D.C., *et al.* (2012). [Climate](#)  
730 [change impacts in multispecies systems: Drought alters food web size structure in a field experiment](#).  
731 *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 2990–2997.



**Figure 1: Network accumulation curves.** (a) Dissimilarity of interactions between common species and (b) dissimilarity in species composition between aggregated local webs and the binary metaweb of host-parasite interactions. Aggregated local webs were obtained by sequentially and randomly selecting a number of binary local webs 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 links and (d) scaling of connectance with the number of sampled binary and probabilistic local webs. For a better comparison with binary webs, local probabilistic webs were derived from the probabilistic metaweb with a false positive and false negative rate of zero. A specific value of  $p$  (the local probability of interaction among potentially interacting species) was used for all local webs within a particular curve. Aggregated probabilistic local webs were obtained by sequentially and randomly selecting a number of probabilistic local webs and aggregating both their species and interactions (with the value of  $p$  adjusting according to eq. 5).



**Figure 2: Spatial scaling of interactions.** Expected number of host-parasite interactions in a network aggregating all probabilistic (a) local and (b) regional 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 the probabilistic metaweb by setting the value of  $p$  (the local probability of interaction among potentially interacting species) to 1, ensuring a conservative comparison between aggregated local webs and metawebs. Aggregated local webs were obtained by aggregating both the species and interactions found within a particular latitudinal window, with the values of  $p$  remaining at their maximum value of 1 following eq. 5.

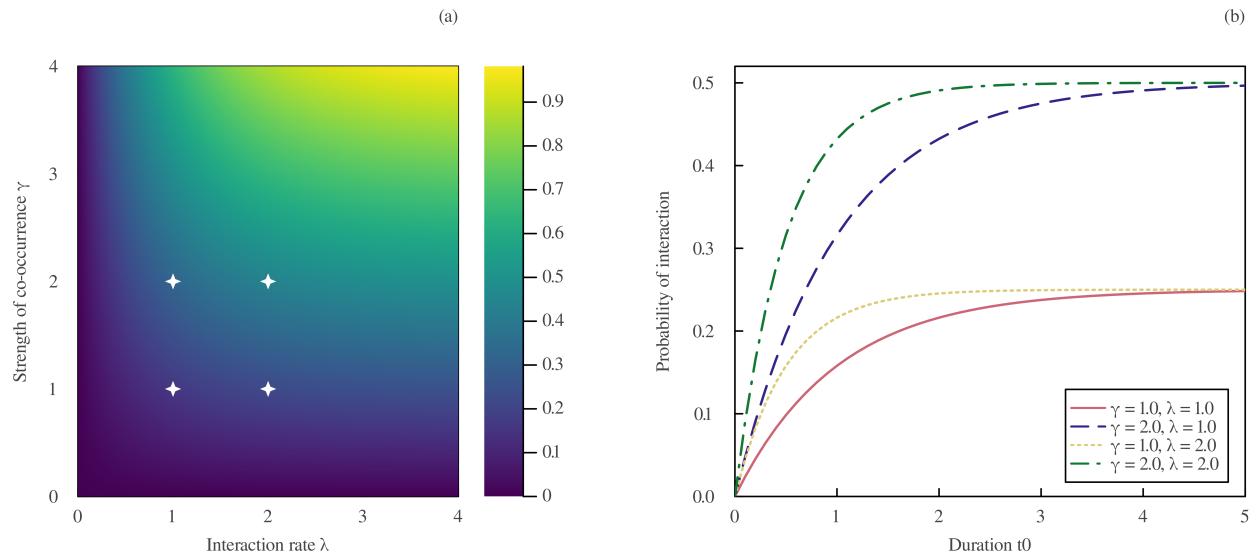
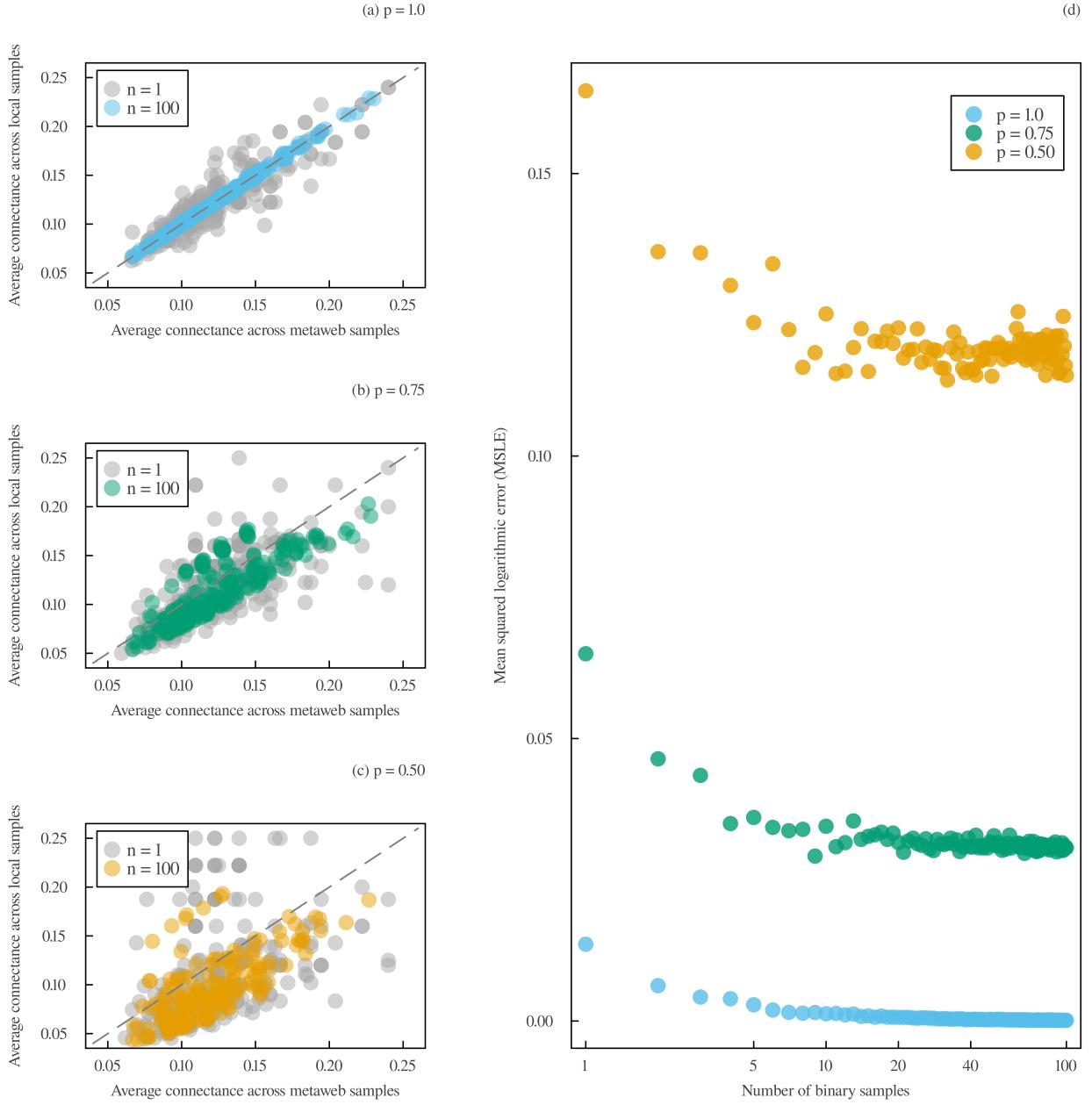


Figure 3: **Parameters of the spatiotemporally explicit model of interactions.** (a) Probability of local interaction given by the process model (eq. 11) under different values of  $\lambda$  and  $\gamma$ , with  $t_0 = 1$ . The parameter values used in the right panel are denoted by the white stars. (b) Scaling of the probability of interaction with the duration parameter  $t_0$  in eq. 11, for different values of  $\lambda$  and  $\gamma$ . In both panels, the individual probabilities of occurrence  $P_i(x, y, z)$  and  $P_j(x, y, z)$  are set to a constant value of 0.5.



**Figure 4: Connectance of sampled binary webs.** Comparison between the average connectance of binary network samples obtained from the probabilistic local and metawebs. Each dot corresponds to a different site. The local probability of interaction between potentially interacting species was set to three different values: (a)  $p = 1.0$ , (b)  $p = 0.75$ , and (c)  $p = 0.50$ . Grey dots represent the outcome of a single trial, while colored dots represent the average connectance of each network across 100 trials. (d) Reduction in the mean squared logarithmic error between the average connectance of binary networks obtained from these two sampling methods as the number of trials increases, for the same values of  $p$  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 metaweb, and then propagating this result to all local webs that include the species potentially engaged in the interactions. Local binary webs were generated by independently sampling binary interactions for each local probabilistic web.