

# 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 local probabilistic networks 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 networks from subsets of metawebs. To support our arguments, we build a spatiotemporally explicit model of probabilistic interactions and develop different case studies using 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 6, 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 is a substantial challenge.  
5 There are methodological and biological constraints that hinder our ability to observe all interactions, such as  
6 the spatial and temporal uncoupling of species (Jordano (1987)) and the large number of possible interactions in  
7 a community, the vast majority of which are rare (Jordano (2016)). More generally, a handful of conditions  
8 must be satisfied for an interaction to be observed locally. First, both species must have overlapping geographic  
9 ranges, i.e. they must co-occur within the region of interest (Blanchet *et al.* (2020)). Second, they must have  
10 some probability of meeting. Probabilities of interspecific encounters are typically low, especially for rare  
11 species with low relative abundances (Canard *et al.* (2012)). The probability that species meet each other also  
12 depends on their biological characteristics, such as the synchronization of their phenology (Olesen *et al.* (2010),  
13 Singer & McBride (2012)) and their discoverability (e.g., Broom & Ruxton (2005)). Finally, when species do  
14 come into contact, an interaction occurs only if their traits are locally compatible  
15 (Poisot *et al.* (2015)), including but not limited to their body phenotypes (Bolnick *et al.* (2011), Stouffer *et al.*  
16 (2011), Gravel *et al.* (2013)) and behavioral choices (Pulliam (1974), Choh *et al.* (2012)). Interactions may also  
17 be influenced by the presence or prevalence of a third species (e.g., of a more profitable prey species) (Golubski  
18 & Abrams (2011), Sanders & van Veen (2012)). Documenting the location and timing of interactions becomes  
19 even more difficult when accounting for the spatiotemporal variability of ecological interactions (Poisot *et al.*  
20 (2012), Poisot *et al.* (2015)). Environmental factors, such as temperature (Angilletta *et al.* (2004)), drought  
21 (Woodward *et al.* (2012)), climate change (Gilman *et al.* (2010), Woodward *et al.* (2010), Araujo *et al.* (2011)),  
22 and habitat modifications (Tylianakis *et al.* (2007)), contribute to this spatiotemporal variability of interactions  
23 by impacting species abundance and traits. In this context, it is unsurprising that our knowledge of ecological  
24 interactions remains limited (Hortal *et al.* (2015)) despite extensive biodiversity data collection (Schmeller *et*  
25 *al.* (2015)).

26 Knowing the biological capacity of two species to interact is necessary but not sufficient for inferring their  
27 interaction at a specific time and space. The recognition of the intrinsic variability of species interactions has  
28 led ecologists to expand their representation of ecological networks to include a probabilistic view of  
29 interactions (Poisot *et al.* (2016), Dallas *et al.* (2017), Fu *et al.* (2021)). This different perspective allows us to

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

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

60 *al.* (2016), Pomeranz *et al.* (2019)) for inference. Topological null models, which generate probabilistic  
61 networks by preserving chosen characteristics of the binary adjacency matrix while intentionally omitting  
62 others (Bascompte *et al.* (2003), Fortuna & Bascompte (2006)), serve as other examples of common  
63 probabilistic network models. Null models can be used to produce underlying distributions of network measures  
64 for null hypothesis significance testing. Many measures have been developed to describe the structure (Poisot *et*  
65 *al.* (2016)) and diversity (Ohlmann *et al.* (2019), Godsoe *et al.* (2022)) of probabilistic networks. These models  
66 and measures support the use of this approach for the study of a wide range of ecological questions, from  
67 making better predictions of species distribution (Cazelles *et al.* (2016)) to forecasting the impact of climate  
68 change on ecological networks (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 networks that necessitate distinct  
75 approaches when applied to key ecological questions: local networks of realized interactions, and regional  
76 networks (metawebs) of potential interactions. We highlight the distinctions in the ecological meaning of these  
77 two representations and show that they yield different statistical outcomes regarding e.g. the spatial and  
78 temporal scaling of interactions and the prediction of binary networks across space. Moreover, there is currently  
79 no metadata standard that could guide the documentation of all types of probabilistic interactions (although see  
80 e.g., Salim *et al.* (2022) who discuss data standards for deterministic mutualistic networks). Well-defined  
81 metadata for probabilistic networks would support more adequate manipulation and integration of interaction  
82 data from different sources and guard against the potential misinterpretations arising from ambiguous  
83 definitions of probabilistic networks. These metadata should outline the nature (i.e., potential or local) and type  
84 (e.g., predatory or pollination) of the interactions, provide information regarding the taxonomic level, identities,  
85 and characteristics (e.g., life stages) of the individuals involved in an interaction, present the mathematical  
86 formulation of probabilities, including clearly identified conditional variables, and describe the methods and  
87 contexts (e.g., location, time, environmental conditions) in which interactions were estimated. Inadequately  
88 documented probabilistic interaction data should be used with caution when analyzing ecological networks.  
89 Although our contribution focuses on food webs, our observations and advice can be applied to other types of

90 ecological networks, from plant-pollinator to host-virus networks. Indeed, excluding networks of indirect  
91 interactions such as competition and facilitation networks (Kéfi *et al.* (2015), Kéfi *et al.* (2016)), most  
92 ecological networks describe probabilities of direct interactions, which are conceptually and mathematically  
93 analogous regardless of their 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 interaction is defined as the biological  
97 capacity of two taxa to interact (i.e., the probability that they *can* theoretically interact) whereas a realized  
98 interaction is the materialization or observation of this interaction in a well-defined space and time (i.e., the  
99 probability that they interact locally). Here, we use the terms *metaweb* (Dunne (2006)) to designate networks of  
100 potential interactions and *local networks* (Poisot *et al.* (2012)) for those of realized interactions. Metawebs are  
101 the network analogues of the species pool, where local networks originate from a subset of both species (nodes)  
102 and interactions (edges) of the regional metaweb (Saravia *et al.* (2022)). Without clear documentation, it can be  
103 challenging to know if published probabilistic networks describe potential or realized interactions, or if  
104 so-called probabilities are in reality *interaction scores* (i.e., non-probabilistic quantitative interactions). When  
105 probabilistic potential interactions are used and interpreted as local interactions (and conversely), this may  
106 generate misleading findings during data analysis. We believe that a better understanding of the differences,  
107 similarities, and relationships between these two probabilistic representations of ecological networks would  
108 alleviate interpretation errors and facilitate a more adequate utilization of interaction data.

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

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

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

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

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

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

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

171 We define space as the set of geographic coordinates  $(x, y, z)$  of the spatial boundaries delineating the system,  
172 whether sampled or targeted. Ecological interactions may exhibit variations spatially and along altitudinal  
173 gradients, as evidenced by changes in hummingbird-plant interactions (Weinstein & Graham (2017b) and  
174 Weinstein & Graham (2017a)) and mosquitoes biting rates (e.g., Kulkarni *et al.* (2006)) at different elevations.  
175 In contrast, time is defined as the specific time period within which interactions were either observed or

predicted. With these definitions in mind, space and time can be conceptualized as distinct patches or time segments. Treating them as discrete dimensions aligns with the common sampling methods of ecological networks and provides actual probabilities of interactions across space and time. This differs from the approach of treating them as continuous variables, where edge values represent probability densities (i.e., relative likelihoods of interactions occurring at infinitesimal locations and instants in time) rather than discrete probability values (which can be obtained by integrating probability densities). By employing discrete locations and time periods, we can quantify both an area  $A$  and a duration  $t$ , which can be readily used in spatiotemporal analyses of ecological networks. For example, 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 locally can be conditional on many environmental and biological factors. One of these is their co-occurrence  $C_{i,j}$ , which is usually a Boolean describing if the geographic distributions of both taxa overlap within the study area. Co-occurrence can be modeled probabilistically, in 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 co-occurrence  $P_{i,j}(x, y, z)$  can be estimated through the application of joint species distribution models (e.g., Pollock *et al.* (2014)). Considering that the probability that two non co-occurring taxa interact locally is zero (i.e.,  $P_N(i \rightarrow j|C = 0) = 0$ ), the probability of local interaction can be obtained by multiplying the probability of interaction given co-occurrence with the probability of co-occurrence:

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

Other important factors that can impact interaction probabilities at the local scale are taxa relative abundance (Canard *et al.* (2012)) and traits (Poisot *et al.* (2015)), as well as environmental factors such as temperature (Angilletta *et al.* (2004)), precipitation (Woodward *et al.* (2012)), habitat structure (Klecka & Boukal (2014)), and presence of other interacting taxa in the network (Pilosof *et al.* (2017), Kéfi *et al.* (2012)). Here, we use the variable  $\Omega$  to describe the biological and ecological context in which interaction probabilities were estimated. For example, if a research team conducts a mesocosm experiment to estimate interaction probabilities between predators and prey with and without shelters,  $\Omega$  would represent the presence or absence of these shelters. Like co-occurrence,  $\Omega$  can also be modeled probabilistically when the stochasticity or uncertainty of environmental 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 network  $N$  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 network  $N$  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 networks. 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
$P_M(i \rightarrow j)$	potential	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. Potential 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 networks, 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 networks, 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  
 235 one-on-one interactions, the extent of sampling effort also contributes to shaping our assessment of metaweb  
 236 properties, as sampling a larger number of local networks allows us to capture more potential interactions  
 237 (McLeod *et al.* (2021a)). A probability of potential interaction in a metaweb  $M$  can be expressed as

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

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

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

242 Every pair of taxa that have confidently been observed to interact at least once can be given a probability of 1  
243 (i.e.,  $P_M(i \rightarrow j) = 1$ ) since we know that they *can* interact. This is not the case in local probabilistic networks,  
244 in which interaction events usually remain stochastic (i.e.,  $P_N(i \rightarrow j) < 1$ ) after empirically observing  
245 interactions because of their intrinsic spatiotemporal variability. Similarly, although rare interactions typically  
246 have low probabilities in local networks, they can have high probabilities in metawebs when the traits of both  
247 taxa are congruent. On the other hand, interactions that were never observed can have low probability values in  
248 both metawebs and local networks, going as low as 0 for forbidden links. However, because of observation  
249 errors due to taxonomic misidentifications and ecological misinterpretations (e.g., due to cryptic species and  
250 interactions, Pringle & Hutchinson (2020)), many observations of interactions are in reality false positives.

251 Likewise, forbidden interactions can be false negatives in metawebs, e.g. if they have been assessed for specific  
252 phenotypes, locations or time. Employing Bayesian models, whether they are mechanistic or  
253 phenomenological, has the potential to enhance the accuracy of our estimation of interaction probabilities  
254 within both local and potential networks (e.g., Bartomeus *et al.* (2016), Cirtwill *et al.* (2019)). This  
255 improvement is achieved in potential networks by leveraging prior information regarding the feasibility of  
256 interactions along with the empirical data on observed interactions.

## 257 Properties of probabilistic networks

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

263 of local networks and metawebs through the presentation of four common applications of probabilistic  
264 interactions. All code and data to reproduce these analyses are available at the Open Science Framework (TBD).

265 **Host-parasite network data**

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

278 [Figure 1 about here.]

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

290 To build probabilistic local networks, we first recognize that local interactions must initially be biologically  
 291 feasible before occurring at a specific time and space. A local probability of interaction  $P_N(i \rightarrow j)$  can be  
 292 expressed as 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 potential  
 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 networks from the binary ones by using the probabilistic metawebs and a  
 296 constant value of  $p$  across interactions. Lower values of  $p$  indicate that feasible interactions rarely occur locally,  
 297 intermediate values around 50% suggest considerable spatiotemporal variability, while higher values indicate  
 298 that potential interactions are nearly always realized. Following eq. 4, the local probabilities of interaction  
 299 between a given pair of taxa consistently remained equal to or below their probability of potential 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 networks, according to different values of  $p$ . When aggregating local  
 303 probabilistic networks, the constancy of the probability of potential interaction across the entire study area  
 304 means that any rise in the probability of local interaction is solely attributable to an increase in  $p$ . The  
 305 probability  $p_{1,2}$  of local interaction among potentially interacting species in an aggregated network  $N_{1,2}$  is  
 306 obtained by:

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

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

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

313 **Taxonomic scaling of interactions**

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

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

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

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

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

## 355 Spatial and temporal scaling of interactions

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

368 quantify the spatiotemporal variability of interactions (eq. 2), they serve as valuable tools for developing and  
369 testing equations describing the scaling of network structure with space and time.

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

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

395

[Figure 2 about here.]

396 **Box 1: A spatiotemporally explicit model of interactions**

397 Predicting local networks across time and space is a pivotal goal of network ecology (Strydom *et al.* (2021)). In  
398 a context of scarcity of interaction data (Jordano (2016)), ecologists must resort to predictive models for  
399 reconstructing networks at fine spatial and temporal scales. For instance, real-time biomonitoring data coupled  
400 with appropriate numerical models (Bohan *et al.* (2017)) can be employed to reconstruct local ecological  
401 networks, opening avenues for in-depth studies on local ecosystem functioning and dynamics. Apart from their  
402 predictive applications, statistical models can also be crafted for descriptive purposes, elucidating key  
403 parameters of interest such as probabilities of interactions. In such instances, the parameter values themselves  
404 offer valuable ecological insights. Various types of models (e.g., Bayesian and machine learning models) have  
405 been used for both predictive and descriptive purposes in the realm of ecological interactions (Strydom *et al.*  
406 (2021)). The probabilistic representation of interactions acknowledges the inherent uncertainty in these models,  
407 typically expressed through probability distributions. We introduce and develop a simple generative mechanistic  
408 model for probabilistic local interactions that takes into consideration their spatiotemporal variability (i.e. a  
409 spatiotemporally explicit model of local interactions). It is essential to note that our model is not designed for  
410 potential interactions, which are scale-independent. Rather, it could prove valuable for predicting local  
411 interactions across time and space by generating new interaction data following parameter inference.

412 As indicated by eq. 1, the probability that two taxa  $i$  and  $j$  interact at a specific location  $(x, y, z)$  is determined by  
413 the product of their probability of interaction given co-occurrence and their probability of co-occurrence. First,  
414 their co-occurrence probability can be calculated using their individual occurrence probabilities  $P_i(x, y, z)$  and  
415  $P_j(x, y, z)$ , along with the strength of association  $\gamma$  between their occurrences and co-occurrence (Cazelles *et al.*  
416 (2016)):

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

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

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

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

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

426 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  
 427 instance, if the duration  $t_0$  is measured in months,  $\lambda$  denote the expected number of interactions per month.  
 428 The occurrence of an interaction between  $i$  and  $j$  can be modeled as a Bernoulli trial with a probability of  
 429  $P_N(i \rightarrow j)$ . Consequently, a Bayesian inference model can be built based on the preceding equations to estimate  
 430 the value of the  $\lambda$  and  $\gamma$  parameters and generate novel interaction data:

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

431

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

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

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

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

### 438 Sampling binary networks

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

451 There are at least two distinct approaches to sample binary networks from probabilistic webs across space, for  
452 example, when attempting to predict a binary network for each of a number of locations within a given region.

453 The first approach involves performing a singular Bernoulli trial for each pair of taxa within the region of  
454 interest based on their regional probability of potential interaction. In employing this approach, every pair of  
455 taxa predicted to interact in the binary metaweb realization will be treated as interacting in all local networks  
456 where they co-occur. This will result in local pairwise interactions without spatial variation. The second  
457 approach is to independently sample each of the local probabilistic networks, which can be achieved by first  
458 generating distinct probabilistic networks for each location. These local probabilistic networks may vary in taxa  
459 composition and interaction probabilities. Subsequently, binary networks can be independently sampled for  
460 each location, introducing spatial variation in binary interactions.

461 In fig. 4, we compare the average connectance of binary networks resulting from these two sampling techniques,  
462 where potential and local interactions are drawn from our host-parasite probabilistic networks, generating a  
463 number of binary network realizations for each site in the dataset. These two sampling approaches yield  
464 different outcomes, particularly for lower values of  $p$ , which denote instances when potential interactions do not

465 consistently manifest locally. Small discrepancies are also apparent between these techniques when we equate  
466 the probability of local interaction to the probability of potential interaction (i.e., when using  $p = 1.0$  in eq. 4),  
467 especially when the number of binary network samples for each location is low. As anticipated, we observe that  
468 sampling binary interactions from the metaweb tends to overestimate connectance on average compared to  
469 sampling them from local networks. Furthermore, we observe an increase in the variability of connectance  
470 when employing a single sample, representing what we consider as a more tangible process leading to the  
471 realization of local and potential interactions in nature.

472 [Figure 4 about here.]

473 The choice of a sampling approach can influence the selection of grid cell size when delineating local  
474 communities within a broader region of interest. In the first approach, pairwise interactions remain constant  
475 irrespective of cell size since they are sampled only once from the regional network. However, in the second  
476 approach, local interaction probabilities are contingent on the network area. For instance, consider networks  $N_1$   
477 and  $N_2$  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  
478  $N_2$  as independent, the probability of interaction between taxa  $i$  and  $j$  in  $N_0$  is given by:

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

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

### 483 Prediction of local networks from metawebs

484 Metawebs serve as a valuable source of ecological information for predicting local networks across time and  
485 space. Binary local networks can be reconstructed by selecting a subset of taxa and interactions from the  
486 metaweb (Dunne (2006)). This implies that metawebs consistently contain more interactions than their  
487 corresponding local networks. Determining the list of taxa to select can be achieved empirically or through  
488 methods like range maps or species distribution models. As the species composition of a community is arguably  
489 less difficult to sample or predict than its interactions, the primary challenge lies in deciding which interactions

490 to select from the metaweb, a task that may necessitate advanced statistical models and ecological expertise.  
491 Inferring the structure of local networks from the metaweb before predicting specific local pairwise interactions  
492 could hold promise (Strydom *et al.* (2021)), considering that the structure of local networks is constrained by  
493 the metaweb (Saravia *et al.* (2022)).

494 Building probabilistic local networks from a probabilistic metaweb involves a reduction in the value of pairwise  
495 interaction probabilities. This decrease is attributed to the prerequisite that two taxa must initially possess the  
496 capacity to interact before engaging in local interactions (eq. 4). Therefore, inferring local networks from their  
497 metaweb while maintaining identical interaction probability values would introduce systematic biases into the  
498 predictions. In such cases, these networks would essentially represent smaller-scale metawebs of potential  
499 interactions, potentially leading to misinterpretations by being perceived as local interactions. As proposed by  
500 McLeod *et al.* (2021b), although metawebs lack the ability to capture the spatiotemporal variability of  
501 interactions, they establish an upper limit for local interactions. In other words, the probability of two taxa  
502 interacting at a specific location and time is consistently lower or equal to the probability of their potential  
503 interaction:

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

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

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

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

513 **Conclusion**

514 In this contribution, we underline the importance of network metadata for adequately interpreting and  
515 manipulating probabilistic interaction data. The mathematical representation of probabilities and their statistical  
516 properties depend on the type of interactions (local or potential) and the conditions under which these  
517 interactions were evaluated. We showed that probabilistic local networks and metawebs differ in their  
518 relationship to spatial and temporal scales, with potential interactions remaining consistent across scales. In  
519 contrast with metawebs, local interactions are measured in a specific context (e.g., in a given area, time, and  
520 biological and environmental conditions) and depend on taxa co-occurrence. These important conceptual  
521 differences bring to light the need to use probabilistic data with caution, for instance when generating binary  
522 network realizations across space and predicting local networks from metawebs. Clear metadata describing the  
523 type of interaction and the variables used in their estimation are required to ensure adequate data manipulation.  
524 Better data practices and foundations for probabilistic thinking in network ecology could enable more reliable  
525 assessments of the spatiotemporal variability and uncertainty of biotic interactions.

526 It is essential to enhance our comprehension of both potential and local interactions, especially considering the  
527 current scarcity of interaction data. However, while sampling biological communities does decrease the  
528 uncertainty of interactions by accumulating evidence for their feasibility and local realization, there is a limit to  
529 how much we can diminish their inherent randomness. In metawebs, probabilities reflect our limited knowledge  
530 of interactions (i.e., our degree of belief that interactions are feasible), which is expected to improve with a  
531 larger volume of data. We should anticipate that potential interactions will become more definitive (with  
532 probabilities approaching 0 or 1) as we investigate various conditions, including different combinations of  
533 species traits. Conversely, in the case of local networks, which can be seen as random instances of metawebs,  
534 randomness cannot be as reduced. In local interactions, probabilities can be divided into a component  
535 representing uncertainty and another representing spatiotemporal variability. Owing to environmental  
536 heterogeneity, there will invariably be instances in which an interaction occurs and others in which it does not,  
537 across different times and locations, irrespective of the extent to which we can improve our knowledge of its  
538 biological feasibility and the local conditions that facilitate its occurrence. When local networks depict  
539 probabilities of observing interactions rather than just their actual occurrence, we must also consider the  
540 observation variability as an additional source of randomness. Every ecological process is stochastic but there is  
541 also a possibility that a phenomenon goes undetected. Quantifying and partitioning this stochasticity will

542 enable us to make more accurate predictions about ecological interactions at various spatial and temporal  
543 scales. This will prove to be of vital importance as our time to understand nature runs out, especially at the  
544 places where the climate change impacts hit harder.

## 545 Acknowledgment

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

## 553 References

- 554 Abrego, N., Roslin, T., Huotari, T., Ji, Y., Schmidt, N.M., Wang, J., *et al.* (2021). [Accounting for species](#)  
555 [interactions is necessary for predicting how arctic arthropod communities respond to climate change.](#)  
556 *Ecography*, 44, 885–896.
- 557 Andrade-Pacheco, R., Rerolle, F., Lemoine, J., Hernandez, L., Meïté, A., Juziwelo, L., *et al.* (2020). [Finding](#)  
558 [hotspots: Development of an adaptive spatial sampling approach.](#) *Scientific Reports*, 10, 10939.
- 559 Angilletta, M.J., Jr., Steury, T.D. & Sears, M.W. (2004). [Temperature, Growth Rate, and Body Size in](#)  
560 [Ectotherms: Fitting Pieces of a Life-History Puzzle1.](#) *Integrative and Comparative Biology*, 44, 498–509.
- 561 Araujo, M.B., Rozenfeld, A., Rahbek, C. & Marquet, P.A. (2011). [Using species co-occurrence networks to](#)  
562 [assess the impacts of climate change.](#) *Ecography*, 34, 897–908.
- 563 Bartomeus, I., Gravel, D., Tylianakis, J.M., Aizen, M.A., Dickie, I.A. & Bernard-Verdier, M. (2016). [A](#)  
564 [common framework for identifying linkage rules across different types of interactions.](#) *Functional Ecology*,  
565 30, 1894–1903.

- 566 Bascompte, J., Jordano, P., Melián, C.J. & Olesen, J.M. (2003). The nested assembly of plant-animal mutualistic  
567 networks. *Proceedings of the National Academy of Sciences*, 100, 9383–9387.
- 568 Berlow, E.L., Neutel, A.-M., Cohen, J.E., De Ruiter, P.C., Ebenman, B., Emmerson, M., et al. (2004).  
569 Interaction strengths in food webs: Issues and opportunities. *Journal of Animal Ecology*, 73, 585–598.
- 570 Blanchet, F.G., Cazelles, K. & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions.  
571 *Ecology Letters*, 23, 1050–1063.
- 572 Bohan, D.A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A.J. & Woodward, G. (2017).  
573 Next-Generation Global Biomonitoring: Large-scale, Automated Reconstruction of Ecological Networks.  
574 *Trends in Ecology & Evolution*, 32, 477–487.
- 575 Bolnick, D.I., Amarasekare, P., Araújo, M.S., Bürger, R., Levine, J.M., Novak, M., et al. (2011). Why  
576 intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution*, 26, 183–192.
- 577 Borrett, S.R. & Scharler, U.M. (2019). Walk partitions of flow in Ecological Network Analysis: Review and  
578 synthesis of methods and indicators. *Ecological Indicators*, 106, 105451.
- 579 Boulangeat, I., Gravel, D. & Thuiller, W. (2012). Accounting for dispersal and biotic interactions to disentangle  
580 the drivers of species distributions and their abundances. *Ecology Letters*, 15, 584–593.
- 581 Brimacombe, C., Bodner, K., Michalska-Smith, M., Poisot, T. & Fortin, M.-J. (2023). Shortcomings of reusing  
582 species interaction networks created by different sets of researchers. *PLOS Biology*, 21, e3002068.
- 583 Broom, M. & Ruxton, G.D. (2005). You can run or you can hide: Optimal strategies for cryptic prey against  
584 pursuit predators. *Behavioral Ecology*, 16, 534–540.
- 585 Brose, U., Ostling, A., Harrison, K. & Martinez, N.D. (2004). Unified spatial scaling of species and their  
586 trophic interactions. *Nature*, 428, 167–171.
- 587 Canard, E., Mouquet, N., Marescot, L., Gaston, K.J., Gravel, D. & Mouillot, D. (2012). Emergence of  
588 Structural Patterns in Neutral Trophic Networks. *PLOS ONE*, 7, e38295.
- 589 Catchen, M.D., Poisot, T., Pollock, L.J. & Gonzalez, A. (2023). The missing link: Discerning true from false  
590 negatives when sampling species interaction networks.
- 591 Cazelles, K., Araujo, M.B., Mouquet, N. & Gravel, D. (2016). A theory for species co-occurrence in interaction  
592 networks. *Theoretical Ecology*, 9, 39–48.

- 593 Chagnon, P.-L. (2015). [Characterizing topology of ecological networks along gradients: The limits of metrics' standardization](#). *Ecological Complexity*, 22, 36–39.
- 594
- 595 Choh, Y., Ignacio, M., Sabelis, M.W. & Janssen, A. (2012). [Predator-prey role reversals, juvenile experience and adult antipredator behaviour](#). *Scientific Reports*, 2, 728.
- 596
- 597 Cirtwill, A.R., Eklöf, A., Roslin, T., Wootton, K. & Gravel, D. (2019). [A quantitative framework for investigating the reliability of empirical network construction](#). *Methods in Ecology and Evolution*, 10, 599 902–911.
- 600 Dallas, T., Park, A.W. & Drake, J.M. (2017). [Predicting cryptic links in host-parasite networks](#). *PLOS Computational Biology*, 13, e1005557.
- 601
- 602 Dunne, J.A. (2006). The Network Structure of Food Webs. In: *Ecological networks: Linking structure and dynamics* (eds. Dunne, J.A. & Pascual, M.). Oxford University Press, pp. 27–86.
- 603
- 604 Elmasri, M., Farrell, M.J., Davies, T.J. & Stephens, D.A. (2020). [A hierarchical Bayesian model for predicting ecological interactions using scaled evolutionary relationships](#). *The Annals of Applied Statistics*, 14, 605 221–240.
- 606
- 607 Elton, C.S. (2001). *Animal Ecology*. University of Chicago Press, Chicago, IL.
- 608 Emmerson, M.C. & Raffaelli, D. (2004). [Predatorprey body size, interaction strength and the stability of a real food web](#). *Journal of Animal Ecology*, 73, 399–409.
- 609
- 610 Fortuna, M.A. & Bascompte, J. (2006). [Habitat loss and the structure of plantanimal mutualistic networks](#). *Ecology Letters*, 9, 281–286.
- 611
- 612 Fu, X., Seo, E., Clarke, J. & Hutchinson, R.A. (2021). [Link Prediction Under Imperfect Detection: Collaborative Filtering for Ecological Networks](#). *IEEE Transactions on Knowledge and Data Engineering*, 613 33, 3117–3128.
- 614
- 615 Galiana, N., Lurgi, M., Claramunt-López, B., Fortin, M.-J., Leroux, S., Cazelles, K., et al. (2018). [The spatial scaling of species interaction networks](#). *Nature Ecology & Evolution*, 2, 782–790.
- 616
- 617 Gilman, S.E., Urban, M.C., Tewksbury, J., Gilchrist, G.W. & Holt, R.D. (2010). [A framework for community interactions under climate change](#). *Trends in Ecology & Evolution*, 25, 325–331.
- 618
- 619 Godsoe, W., Murray, R. & Iritani, R. (2022). [Species interactions and diversity: A unified framework using Hill numbers](#). *Oikos*, n/a, e09282.
- 620

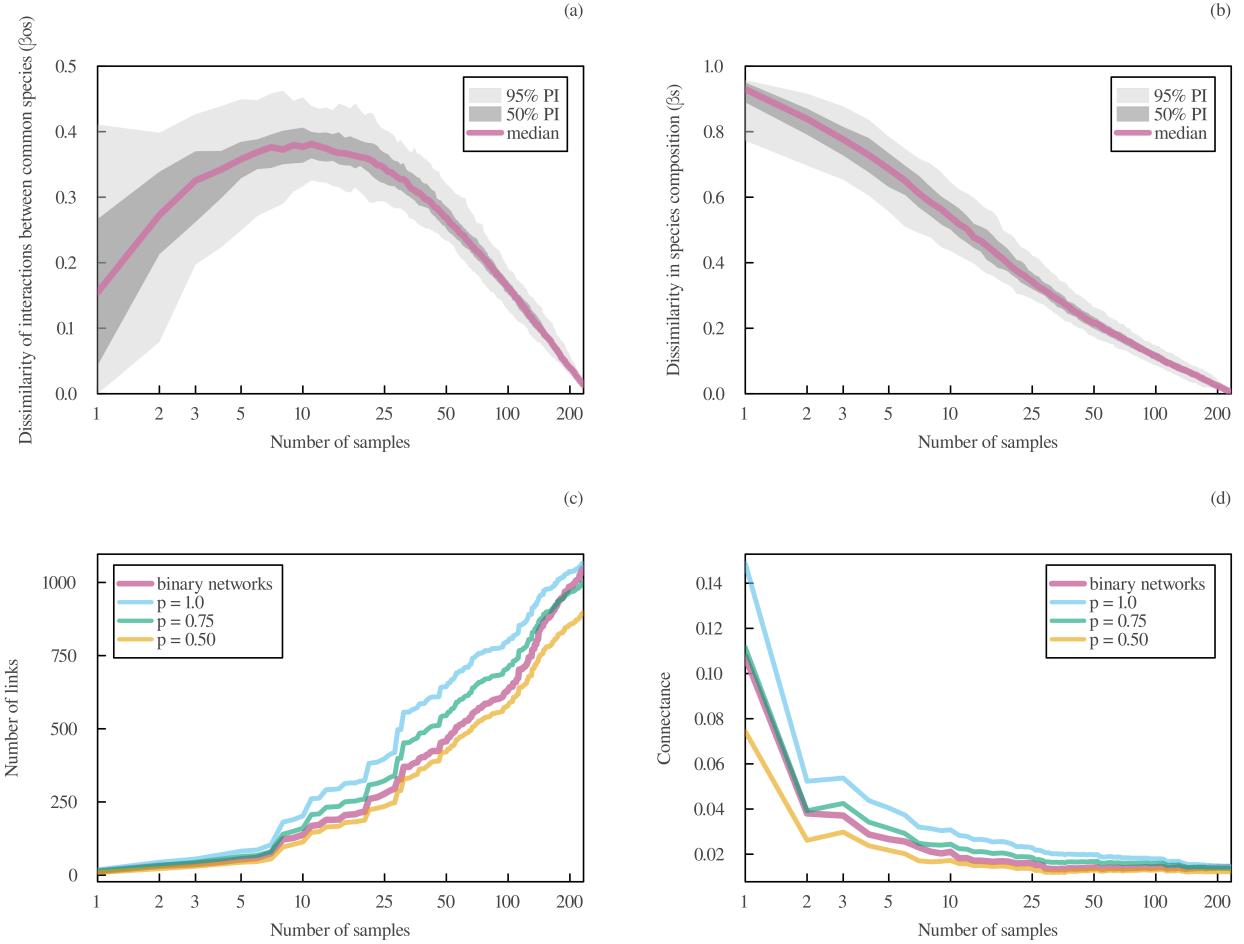
- 621 Golubski, A.J. & Abrams, P.A. (2011). [Modifying modifiers: What happens when interspecific interactions](#)  
622 [interact?](#) *Journal of Animal Ecology*, 80, 1097–1108.
- 623 Gomez, J.M., Verdu, M. & Perfectti, F. (2010). [Ecological interactions are evolutionarily conserved across the](#)  
624 [entire tree of life.](#) *Nature*, 465, 918–U6.
- 625 Gonzalez, A. & Londoño, M.C. (2022). [Monitor biodiversity for action.](#) *Science*, 378, 1147–1147.
- 626 Gonzalez-Varo, J.P. & Traveset, A. (2016). [The Labile Limits of Forbidden Interactions.](#) *Trends in Ecology &*  
627 *Evolution*, 31, 700–710.
- 628 Gravel, D., Baiser, B., Dunne, J.A., Kopelke, J.-P., Martinez, N.D., Nyman, T., *et al.* (2019). [Bringing Elton](#)  
629 [and Grinnell together: A quantitative framework to represent the biogeography of ecological interaction](#)  
630 [networks.](#) *Ecography*, 42, 401–415.
- 631 Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. (2013). [Inferring food web structure from](#)  
632 [predatorprey body size relationships.](#) *Methods in Ecology and Evolution*, 4, 1083–1090.
- 633 Guimarães, P.R. (2020). [The Structure of Ecological Networks Across Levels of Organization.](#) *Annual Review*  
634 *of Ecology, Evolution, and Systematics*, 51, 433–460.
- 635 Guimerà, R. & Sales-Pardo, M. (2009). [Missing and spurious interactions and the reconstruction of complex](#)  
636 [networks.](#) *Proceedings of the National Academy of Sciences*, 106, 22073–22078.
- 637 Hemprich-Bennett, D.R., Oliveira, H.F.M., Le Comber, S.C., Rossiter, S.J. & Clare, E.L. (2021). [Assessing the](#)  
638 [impact of taxon resolution on network structure.](#) *Ecology*, 102, e03256.
- 639 Herrera, C.M. (1989). [Pollinator abundance, morphology, and flower visitation rate: Analysis of the “quantity”](#)  
640 [component in a plant-pollinator system.](#) *Oecologia*, 80, 241–248.
- 641 Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M. & Ladle, R.J. (2015). [Seven Shortfalls](#)  
642 [that Beset Large-Scale Knowledge of Biodiversity.](#) *Annual Review of Ecology, Evolution, and Systematics*,  
643 46, 523–549.
- 644 Jordano, P. (1987). [Patterns of Mutualistic Interactions in Pollination and Seed Dispersal: Connectance,](#)  
645 [Dependence Asymmetries, and Coevolution.](#) *The American Naturalist*, 129, 657–677.
- 646 Jordano, P. (2016). [Sampling networks of ecological interactions.](#) *Functional Ecology*, 30, 1883–1893.

- 647 Jordano, P., Bascompte, J. & Olesen, J.M. (2003). Invariant properties in coevolutionary networks of  
648 plant-animal interactions. *Ecology Letters*, 6, 69–81.
- 649 Kéfi, S., Berlow, E.L., Wieters, E.A., Joppa, L.N., Wood, S.A., Brose, U., et al. (2015). Network structure  
650 beyond food webs: Mapping non-trophic and trophic interactions on Chilean rocky shores. *Ecology*, 96,  
651 291–303.
- 652 Kéfi, S., Berlow, E.L., Wieters, E.A., Navarrete, S.A., Petchey, O.L., Wood, S.A., et al. (2012). More than a  
653 meal... integrating non-feeding interactions into food webs. *Ecology Letters*, 15, 291–300.
- 654 Kéfi, S., Miele, V., Wieters, E.A., Navarrete, S.A. & Berlow, E.L. (2016). How Structured Is the Entangled  
655 Bank? The Surprisingly Simple Organization of Multiplex Ecological Networks Leads to Increased  
656 Persistence and Resilience. *PLOS Biology*, 14, e1002527.
- 657 Klecka, J. & Boukal, D.S. (2014). The effect of habitat structure on prey mortality depends on predator and  
658 prey microhabitat use. *Oecologia*, 176, 183–191.
- 659 Kopalke, J.-P., Nyman, T., Cazelles, K., Gravel, D., Vissault, S. & Roslin, T. (2017). Food-web structure of  
660 willow-galling sawflies and their natural enemies across Europe. *Ecology*, 98, 1730–1730.
- 661 Kulkarni, M.A., Kweka, E., Nyale, E., Lyatuu, E., Mosha, F.W., Chandramohan, D., et al. (2006).  
662 Entomological Evaluation of Malaria Vectors at Different Altitudes in Hai District, Northeastern Tanzania.  
663 *Journal of Medical Entomology*, 43, 580–588.
- 664 McCann, K. (2007). Protecting biostructure. *Nature*, 446, 29–29.
- 665 McCann, K.S. (2011). Food webs (MPB-50). In: *Food Webs (MPB-50)*. Princeton University Press.
- 666 McLeod, A., Leroux, S.J., Gravel, D., Chu, C., Cirtwill, A.R., Fortin, M.-J., et al. (2021a). Sampling and  
667 asymptotic network properties of spatial multi-trophic networks. *Oikos*, 130, 2250–2259.
- 668 McLeod, A., Leroux, S.J., Gravel, D., Chu, C., Cirtwill, A.R., Fortin, M.-J., et al. (2021b). Sampling and  
669 asymptotic network properties of spatial multi-trophic networks. *Oikos*, n/a.
- 670 McLeod, A.M., Leroux, S.J. & Chu, C. (2020). Effects of species traits, motif profiles, and environment on  
671 spatial variation in multi-trophic antagonistic networks. *Ecosphere*, 11, e03018.
- 672 Melián, C.J., Vilas, C., Baldó, F., González-Ortegón, E., Drake, P. & Williams, R.J. (2011). Eco-evolutionary  
673 Dynamics of Individual-Based Food Webs. In: *Advances in Ecological Research*, The Role of Body Size in  
674 Multispecies Systems (ed. Belgrano, A.). Academic Press, pp. 225–268.

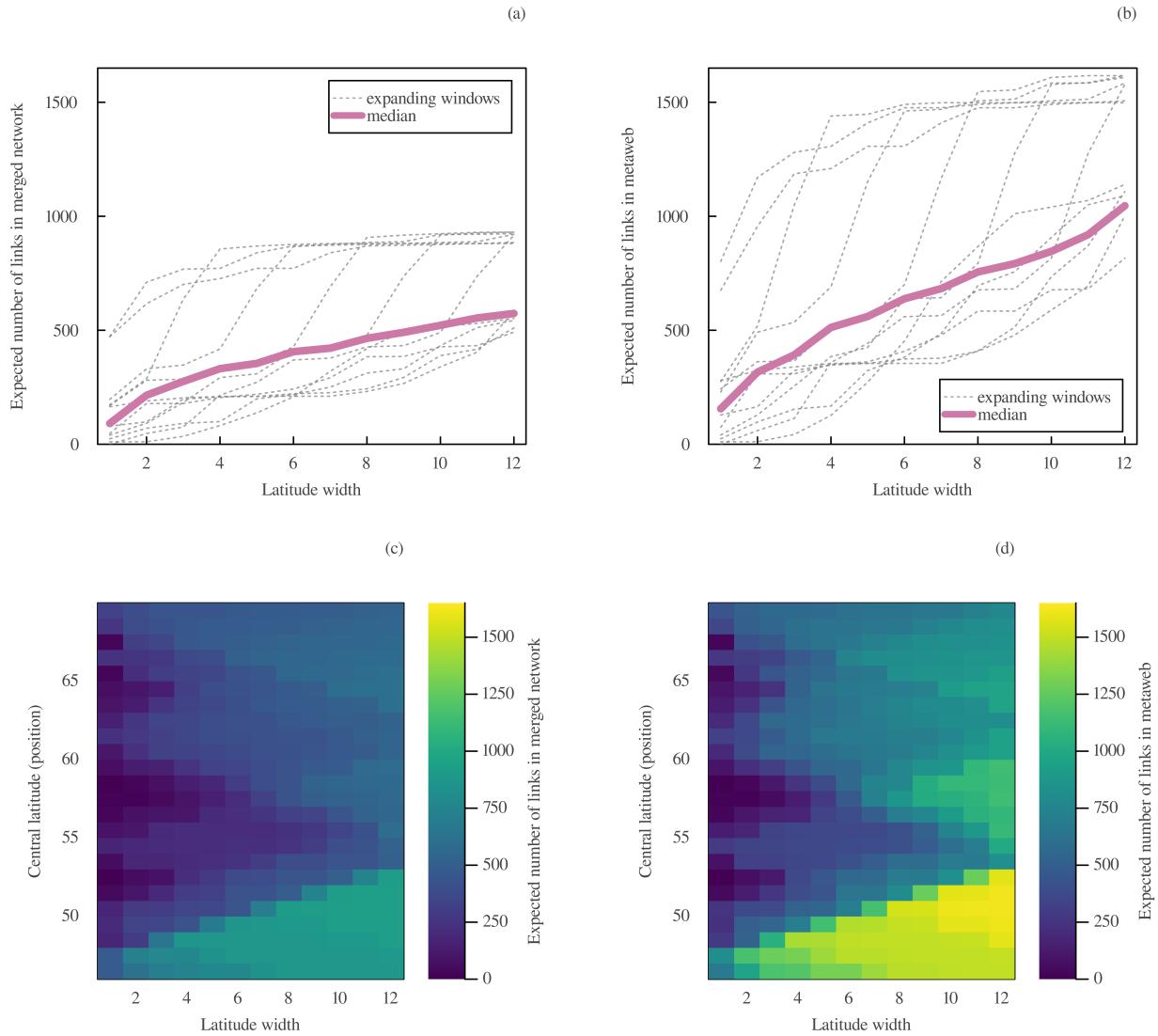
- 675 Ohlmann, M., Miele, V., Dray, S., Chalmandrier, L., O'Connor, L. & Thuiller, W. (2019). [Diversity indices for ecological networks: A unifying framework using Hill numbers](#). *Ecology Letters*, 22, 737–747.
- 677 Olesen, J.M., Bascompte, J., Dupont, Y.L., Elberling, H., Rasmussen, C. & Jordano, P. (2010). [Missing and forbidden links in mutualistic networks](#). *Proceedings of the Royal Society B: Biological Sciences*, 278, 678 725–732.
- 680 Pascual, M., Dunne, J.A. & Dunne, J.A. (2006). *Ecological Networks: Linking Structure to Dynamics in Food Webs*. Oxford University Press, USA.
- 682 Petchey, O.L., Beckerman, A.P., Riede, J.O. & Warren, P.H. (2008). [Size, foraging, and food web structure](#). *Proceedings of the National Academy of Sciences*, 105, 4191–4196.
- 684 Pilosof, S., Porter, M.A., Pascual, M. & Kéfi, S. (2017). [The multilayer nature of ecological networks](#). *Nature Ecology & Evolution*, 1, 1–9.
- 686 Poisot, T., Canard, E., Mouillot, D., Mouquet, N. & Gravel, D. (2012). [The dissimilarity of species interaction networks](#). *Ecology Letters*, 15, 1353–1361.
- 688 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.
- 690 Poisot, T. & Gravel, D. (2014). [When is an ecological network complex? Connectance drives degree distribution and emerging network properties](#). *PeerJ*, 2, e251.
- 692 Poisot, T., Stouffer, D.B. & Gravel, D. (2015). [Beyond species: Why ecological interaction networks vary through space and time](#). *Oikos*, 124, 243–251.
- 694 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 Distribution Model \(JSDM\)](#). *Methods in Ecology and Evolution*, 5, 397–406.
- 696 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.
- 698 Pringle, R.M. & Hutchinson, M.C. (2020). [Resolving Food-Web Structure](#). *Annual Review of Ecology, Evolution, and Systematics*, 51, 55–80.
- 700 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.

- 703 Pulliam, H.R. (1974). [On the Theory of Optimal Diets](#). *The American Naturalist*, 108, 59–74.
- 704 Rooney, N. & McCann, K.S. (2012). [Integrating food web diversity, structure and stability](#). *Trends in Ecology & Evolution*, 27, 40–46.
- 706 Salim, J.A., Saraiva, A.M., Zermoglio, P.F., Agostini, K., Wolowski, M., Drucker, D.P., *et al.* (2022). [Data standardization of plantpollinator interactions](#). *GigaScience*, 11, giac043.
- 708 Sanders, D. & van Veen, F.J.F. (2012). [Indirect commensalism promotes persistence of secondary consumer species](#). *Biology Letters*, 8, 960–963.
- 710 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.
- 712 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.
- 714 Singer, M.C. & McBride, C.S. (2012). [Geographic mosaics of species' association: A definition and an example driven by plantinsect phenological synchrony](#). *Ecology*, 93, 2658–2673.
- 716 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.
- 718 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.
- 720 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.
- 724 Tylianakis, J.M., Tscharntke, T. & Lewis, O.T. (2007). [Habitat modification alters the structure of tropical hostparasitoid food webs](#). *Nature*, 445, 202–205.
- 726 Vázquez, D.P., Peralta, G., Cagnolo, L., Santos, M. & Igual, E. autores contribuyeron por. (2022). [Ecological interaction networks. What we know, what we don't, and why it matters](#). *Ecología Austral*, 32, 670–697.
- 728 Weinstein, B.G. & Graham, C.H. (2017a). [On comparing traits and abundance for predicting species interactions with imperfect detection](#). *Food Webs*, 11, 17–25.

- 730 Weinstein, B.G. & Graham, C.H. (2017b). Persistent bill and corolla matching despite shifting temporal  
731 resources in tropical hummingbird-plant interactions. *Ecology Letters*, 20, 326–335.
- 732 Wood, S.A., Russell, R., Hanson, D., Williams, R.J. & Dunne, J.A. (2015). Effects of spatial scale of sampling  
733 on food web structure. *Ecology and Evolution*, 5, 3769–3782.
- 734 Woodward, G., Benstead, J.P., Beveridge, O.S., Blanchard, J., Brey, T., Brown, L.E., et al. (2010). Chapter 2 -  
735 Ecological Networks in a Changing Climate. In: *Advances in Ecological Research*, Ecological Networks  
736 (ed. Woodward, G.). Academic Press, pp. 71–138.
- 737 Woodward, G., Brown, L.E., Edwards, F.K., Hudson, L.N., Milner, A.M., Reuman, D.C., et al. (2012). Climate  
738 change impacts in multispecies systems: Drought alters food web size structure in a field experiment.  
739 *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 networks and the binary metaweb of host-parasite interactions. Aggregated local networks were obtained by sequentially and randomly selecting a number of binary local networks 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 networks. For a better comparison with binary networks, local probabilistic networks were derived from the probabilistic metaweb with a false positive and false negative rates of zero. A specific value of  $p$  (the local probability of interaction among potentially interacting species) was used for all local networks within a particular curve. Aggregated probabilistic local networks were obtained by sequentially and randomly selecting a number of probabilistic local networks 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) potential 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) potential interactions found in windows of specified length and central latitudes. Probabilities of potential 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 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  $p$  remaining at their maximum value of 1 following eq. 5.

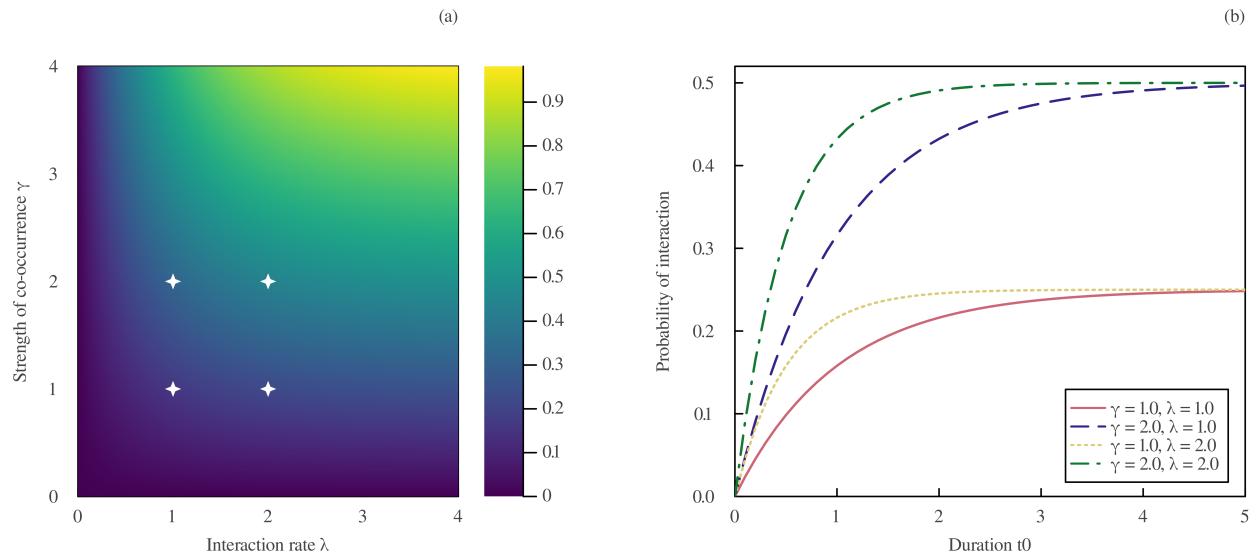
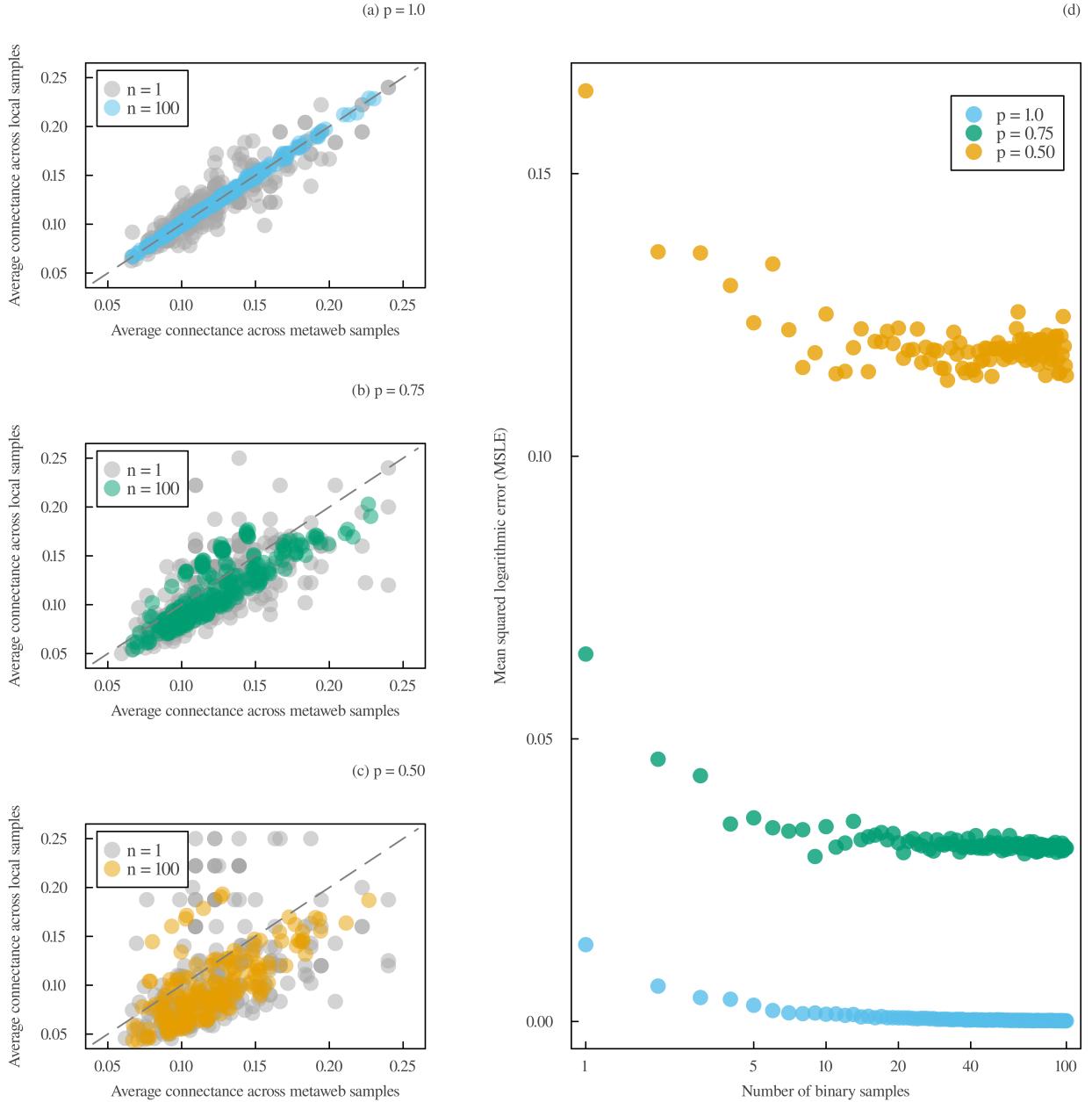


Figure 3: **Parameters of the spatiotemporally explicit model of interactions.** (a) Probability of local interaction given by the process model (eq. 11) under different values of  $\lambda$  and  $\gamma$ , with  $t_0 = 1$ . The parameter values used in the right panel are denoted by the white stars. (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 networks.** Comparison between the average connectance of binary network samples obtained from the probabilistic local networks and metaweb. 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 potential 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 networks that include the species potentially engaged in the interactions. Local binary networks were generated by independently sampling binary interactions for each local probabilistic network.