

The ecological interpretation of probabilistic networks

Francis Banville^{1,2,3} Tanya Strydom^{1,3} Penelope S. A. Blyth⁴ Chris Brimacombe⁵ Michael Catchen^{3,6}
Gabriel Dansereau^{1,3} Dominique Gravel^{2,3} Gracielle Higino² Thomas Malpas⁴ Hana Mayall⁴
Kari Norman¹ Timothée Poisot^{1,3}

¹ Université de Montréal ² Université de Sherbrooke ³ Quebec Centre for Biodiversity Science ⁴ University of Sheffield ⁵ University of Toronto ⁶ 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: *February 11, 2024*

1 Introduction

2 As we try to navigate global biodiversity change, filling in knowledge gaps about biodiversity becomes
3 instrumental to monitoring and mitigating those changes (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 (Polis (1991), Pascual *et al.* (2006)). There are methodological and
6 biological constraints that hinder our ability to observe all ecological interactions, leading to significant
7 uncertainties in our understanding of these interactions. For example, the spatial and temporal uncoupling of
8 species (Jordano (1987)) and the large number of rare and cryptic interactions in a community contribute to
9 these uncertainties (Jordano (2016)). More generally, a handful of conditions must be satisfied for an interaction
10 to be observed locally. First, both species must have overlapping geographic ranges, i.e. they must co-occur
11 within the region of interest (Blanchet *et al.* (2020)). Second, they must have some probability of meeting.
12 Probabilities of interspecific encounters are typically low, especially for rare species with low relative
13 abundances (Canard *et al.* (2012)). The probability that species meet each other also depends on their
14 biological characteristics, such as the synchronization of their phenology (Olesen *et al.* (2010), Singer &
15 McBride (2012)) and their discoverability (e.g., Broom & Ruxton (2005)). Finally, when species do come into
16 contact, an interaction occurs only if their traits are locally compatible
17 (Poisot *et al.* (2015)), including but not limited to their body phenotypes (Bolnick *et al.* (2011), Stouffer *et al.*
18 (2011), Gravel *et al.* (2013)) and behavioral choices (Pulliam (1974), Choh *et al.* (2012)). Interactions may also
19 be influenced by the presence or prevalence of a third species (e.g., of a more profitable prey species) (Golubski
20 & Abrams (2011), Sanders & van Veen (2012)). Documenting the location and timing of interactions becomes
21 even more difficult when accounting for the spatiotemporal variability of ecological interactions (Poisot *et al.*
22 (2012), Poisot *et al.* (2015)). Environmental factors, such as temperature (Angilletta *et al.* (2004)), drought
23 (Woodward *et al.* (2012)), climate change (Gilman *et al.* (2010), Woodward *et al.* (2010), Araujo *et al.* (2011)),
24 and habitat modifications (Tylianakis *et al.* (2007)), contribute to this spatiotemporal variability of interactions
25 by impacting species abundance and traits. Even after satisfying all these conditions, there remains a possibility
26 that the interaction does not occur locally, either due to the intricate nature of the system or simply by chance. If
27 it does occur, it might still go unnoticed, particularly if it happens infrequently. In this context, it is unsurprising
28 that our knowledge of ecological interactions remains limited (Hortal *et al.* (2015)) despite extensive
29 biodiversity data collection (Schmeller *et al.* (2015)).

30 Knowing the biological capacity of two species to interact directly (via e.g., trophic interactions) is necessary
31 but not sufficient for inferring their interaction at a specific time and space. The recognition of the intrinsic
32 variability of species interactions has led ecologists to expand their representation of ecological networks (also
33 known as ecological webs) to include a probabilistic view of interactions (Poisot *et al.* (2016), Dallas *et al.*
34 (2017), Fu *et al.* (2021)). This different perspective allows us to fill in the Eltonian shortfall (Hortal *et al.*
35 (2015)) by modeling the probability of detecting interactions, which can be an important tool for directing
36 efforts and taking action, especially in places where access and resources for research are scarce. Representing
37 interactions probabilistically enables us to capture the spatiotemporal variability of the aforementioned
38 ecological processes and the uncertainties associated with their measurement. As opposed to binary
39 deterministic webs, in which interactions are regarded as either occurring or not, probabilistic webs, within a
40 Bayesian framework, express our degree of belief (or confidence) regarding the occurrence of interactions.
41 Based on the scale at which they are estimated, probabilistic interactions may reflect our level of confidence in
42 whether interactions will be observed, realized, or biologically feasible. As an illustration, we could outline a
43 situation in which there is a 50% certainty that an interaction occurs 50% of the time, or that there is a 50%
44 certainty that it simply occurs. Our level of confidence should be more definitive (approaching either 0 or 1) as
45 we extend our sampling to a broader area and over a longer time period, thereby diminishing the uncertainty of
46 the interactions (but not necessarily the estimation of their variability). In the broadest sense, binary networks
47 are also a type of probabilistic network, in which the numerical value of an interaction is restrained to 0
48 (non-occurring) or 1 (occurring). Yet, for the sake of clarity, we omit binary networks from our discussion of
49 probabilistic networks in this contribution. In probabilistic webs, only forbidden interactions (i.e., interactions
50 prohibited by biological traits or species absence, Jordano *et al.* (2003), Olesen *et al.* (2010)) have a probability
51 value of zero by default, provided that intraspecific trait variability is considered (Gonzalez-Varo & Traveset
52 (2016)).

53 By accounting for the uncertainty of interactions, probabilistic webs may provide a more realistic portrait of
54 species interactions and network structure (i.e. community-level properties), which are major drivers of the
55 functioning, dynamics, and resilience of ecosystems worldwide (Proulx *et al.* (2005), McCann (2007), McCann
56 (2011), Rooney & McCann (2012)). Moreover, the application and development of computational methods in
57 network ecology, often based on a probabilistic representation of interactions, can alleviate (and guide) the
58 sampling efforts required to document species interactions (Strydom *et al.* (2021)). For example, statistical
59 models can be used to estimate the uncertainty of pairwise interactions (Cirtwill *et al.* (2019)) and the

60 probability of missing (false negatives) and spurious (false positives) interactions (Guimerà & Sales-Pardo
61 (2009)). Considering the high rate of false negatives in species interaction data due to the difficulty of
62 witnessing rare interactions (Catchen *et al.* (2023)), these models could inform the identification of priority
63 sampling locations of ecological webs where data collection would yield the most valuable information, thereby
64 reducing errors. Optimization models for sampling locations have mostly found applications in biological
65 systems that are not networks, such as identifying priority sampling sites for disease hotspots (Andrade-Pacheco
66 *et al.* (2020)), but there is substantial promise in applying them to probabilistic ecological interactions.
67 Statistical models can also be used to generate predictions of ecological webs without prior knowledge of
68 pairwise interactions, for instance using body size (Petchey *et al.* (2008), Gravel *et al.* (2013)), phylogeny
69 (Elmasri *et al.* (2020), Strydom *et al.* (2022)), or a combination of niche and neutral processes (Bartomeus *et*
70 *al.* (2016), Pomeranz *et al.* (2019)) for inference. Topological null models, which generate probabilistic
71 networks by preserving chosen characteristics of the binary adjacency matrix while intentionally omitting
72 others (Bascompte *et al.* (2003), Fortuna & Bascompte (2006)), serve as other examples of common
73 probabilistic network models. Null models can be used to produce underlying distributions of network measures
74 for null hypothesis significance testing. Many measures have been developed to describe the structure (Poisot *et*
75 *al.* (2016)) and diversity (Ohlmann *et al.* (2019), Godsoe *et al.* (2022)) of probabilistic webs. These models and
76 measures support the use of this approach for the study of a wide range of ecological questions, from making
77 better predictions of species distribution (Cazelles *et al.* (2016)) to forecasting the impact of climate change on
78 ecological webs (Gilman *et al.* (2010)).

79 The lack of clear guidelines on the use of probabilistic interaction data is worrisome, both for data producers
80 and re-users who generate and manipulate these numbers. This is concerning because sampling strategies and
81 decisions regarding network construction can affect our understanding of network properties (Brimacombe *et*
82 *al.* (2023)). Besides methodological difficulties that may arise when assessing probabilistic interactions, a
83 precise definition of probabilistic interactions appears to be lacking, making the estimation and use of these data
84 more difficult. We aim to take a step back by outlining different ways in which probabilistic interactions are
85 defined and used in network ecology. We distinguish two broad categories of probabilistic webs that necessitate
86 distinct approaches when applied to key ecological questions: local webs describing probabilities of realized
87 interactions, and regional webs (metawebs) describing probabilities of potential interactions. We highlight the
88 distinctions in the ecological meaning of these two representations and show that they yield different statistical
89 outcomes regarding e.g. the spatial and temporal scaling of interactions and the prediction of binary webs

90 across space. Moreover, there is currently no metadata standard that could guide the documentation of all types
91 of probabilistic interactions (although see e.g., Salim *et al.* (2022) who discuss data standards for deterministic
92 mutualistic webs). Well-defined metadata for probabilistic webs would support more adequate manipulation
93 and integration of interaction data from different sources and guard against possible misinterpretations arising
94 from ambiguous definitions of probabilistic networks. These metadata should outline the nature (i.e., local or
95 regional) and type (e.g., predatory or pollination) of the interactions, provide information regarding the
96 taxonomic level, identities, and characteristics (e.g., life stages) of the individuals involved in an interaction,
97 present the mathematical formulation of probabilities, including clearly identified conditional variables (e.g.,
98 spatial and temporal scales), and describe the methods and contexts (e.g., location, time, environmental
99 conditions) in which interactions were estimated. Inadequately documented probabilistic interaction data should
100 be used with caution when analyzing ecological webs. Our observations and advice can be applied to many
101 types of ecological networks, from food webs to host-virus networks. Indeed, excluding networks of indirect
102 interactions such as competition and facilitation networks (Kéfi *et al.* (2015), Kéfi *et al.* (2016)), most
103 ecological webs describe probabilities of direct interactions, which are conceptually and mathematically
104 analogous regardless of their biological type (e.g., trophic and parasitic interactions).

105 **Probabilistic representations of interactions**

106 One of the first aspects to take into consideration when estimating or interpreting probabilities of interactions is
107 knowing if they describe potential or realized interactions. A potential (regional) interaction is defined as the
108 biological capacity of two taxa to interact (i.e., the probability that they *can* theoretically interact) whereas a
109 realized (local) interaction is the materialization or observation of this interaction in a well-defined space and
110 time (i.e., the probability that they interact locally). Here, we use the terms *metaweb* (Dunne (2006)) to
111 designate regional webs of potential interactions and *local webs* (Poisot *et al.* (2012)) for those of realized
112 interactions. Metawebs are the network analogs of the species pool, where local webs originate from a subset of
113 both species (nodes) and interactions (edges) of the regional metaweb (Saravia *et al.* (2022)). Without clear
114 documentation, it can be challenging to know if published probabilistic webs describe local or regional
115 interactions (tbl. 1 provides examples of studies employing both types of probabilistic networks), or if so-called
116 probabilities are in reality *interaction scores* (i.e., non-probabilistic quantitative interactions). When
117 probabilistic regional interactions are used and interpreted as local interactions (and conversely), this may

118 generate misleading findings during data analysis. We believe that a better understanding of the differences,
119 similarities, and relationships between these two probabilistic representations of ecological webs would
120 alleviate interpretation errors and facilitate a more adequate utilization of interaction data.

121 **Pairwise interactions: the building blocks of ecological networks**

122 Local and metawebs, like any type of network, are made of nodes and edges that can be represented at different
123 levels of organization and precision. The basic unit of food webs and other ecological networks are individuals
124 that interact with each other (e.g., by predation, Elton (2001)), forming individual-based networks (Melián *et al.*
125 (2011)). The aggregation of these individuals into more or less homogeneous groups (e.g., populations, species,
126 families) allows us to represent nodes at broader taxonomic scales, which affects our interpretation of the
127 properties of these systems (Guimarães (2020), Hemprich-Bennett *et al.* (2021)).

128 Edges linking nodes can describe a variety of interaction measures. Ecologists have traditionally represented
129 interactions as binary objects that were considered realized after observing at least one individual from group i
130 interact with at least another individual from group j . In a binary adjacency matrix B , the presence or absence
131 of an interaction $B_{i \rightarrow j}$ between two taxa can be viewed as the result of a Bernoulli process

132 $B_{i \rightarrow j} \sim \text{Bernoulli}(P(B_{i \rightarrow j}))$, with $P(B_{i \rightarrow j})$ being the probability of interaction that characterizes our limited
133 knowledge of the system and its intrinsic spatiotemporal variability. In probabilistic networks, $P(B_{i \rightarrow j})$ are edge
134 values, and the only two possible outcomes are the presence ($B_{i \rightarrow j} = 1$) or absence ($B_{i \rightarrow j} = 0$) of an interaction
135 between each pair of nodes. Depending on the type of probabilistic network (local or metaweb), the
136 mathematical formulation and interpretation of stochastic parameters like $P(B_{i \rightarrow j})$ can be linked to
137 environmental and biological factors such as species relative abundance, traits, area, and time (tbl. 1), for
138 example using logistic regression with a logit link function with continuous explanatory variables. Predicting
139 the number of local webs in which the interaction occurs can be achieved by using a Binomial distribution,
140 assuming a constant probability of interaction and independence between networks (trials). When considering
141 uncertainties around the estimation of $P(B_{i \rightarrow j})$, a Beta distribution can also be used to encompass all possible
142 probability values. In that case, a Beta-Binomial distribution can be used to predict the number of networks in
143 which the interaction occurs. Observing an interaction between two taxa at a given location and time provides
144 important information that can be used to update previous estimates of $P(B_{i \rightarrow j})$, informing us on the biological
145 capacity of both taxa to interact and the environmental conditions that enabled them to interact locally.

Even though binary webs constitute a highly valuable source of ecological information (Pascual *et al.* (2006)), they overlook important factors regarding interaction strengths. Represented in a quantitative adjacency matrix W as numbers not confined to the $[0, 1]$ range, interaction strengths better describe the energy flows, demographic impacts or frequencies of interactions between nodes (Berlow *et al.* (2004), Borrett & Scharler (2019)), with $W_{i \rightarrow j}$ being a natural \mathbb{N} or real \mathbb{R} number depending on the measure. For example, they may represent local interaction rates between pairs of taxa (e.g., the flower-visiting rates of pollinators in a mutualistic network, Herrera (1989)). When interaction strengths characterize predation pressure on prey taxa in food webs, they can serve as good estimators of the parameters describing species interactions in a Lotka-Volterra model (e.g., Emmerson & Raffaelli (2004)). The extra amount of ecological information in quantitative networks typically comes at a cost of greater sampling effort and data volume in predictive models (Strydom *et al.* (2021)), which can lead to relatively high levels of uncertainties when inferring quantitative webs with limited data.

Just like binary networks, the uncertainty and spatiotemporal variability of interaction strengths can be represented probabilistically. However, the need to estimate the probability distribution of all possible values of interaction strengths can make the inference of probabilities more challenging in quantitative webs compared to binary webs, which require only one probability estimate for each interaction. Interaction strengths can follow various probability distributions depending on the measure used. For instance, they can follow a Poisson distribution $W_{i \rightarrow j} \sim \text{Poisson}(\lambda_{i \rightarrow j})$ when predicting frequencies of interactions between pairs of nodes, with $\lambda_{i \rightarrow j}$ being the expected rate at which individuals of taxa i and j interact (e.g., the average number of prey j consumed by all predators i in a given time period). The Poisson distribution can also be 0-inflated after initially modeling non-interacting taxa (e.g., Boulangeat *et al.* (2012) employ a 0-inflated model to analyze species abundance following the modeling of species presence and absence), which constitute the majority of taxa pairs in most local webs due to their typically high sparseness (Jordano (2016)). Because of the methodological difficulties typically encountered when building deterministic quantitative webs (which are only partially mitigated by models such as Ecopath, Plagányi & Butterworth (2004)), binary webs, which are easier to sample (Jordano (2016)) and predict (Strydom *et al.* (2021)), have been more frequently studied and modeled. Moreover, most published probabilistic networks and methods describe probabilistic interactions whose outcome is binary (whether interaction probabilities are regarded as constant or variable, e.g. represented by a Beta distribution), which underlines the need for better guidelines regarding the interpretation and manipulation of these types of webs first. For these reasons, our primary focus in this contribution will be on

176 addressing the challenges in interpreting and using interaction probabilities in Bernoulli distributions, in both
177 probabilistic local and metawebs.

178 **Local webs: communities interacting in space and time**

179 Probabilistic local webs describe how likely taxa are to interact at a given location and time period (i.e.,
180 interactions are contingent upon the environmental and biological conditions of the community). In local webs,
181 edges commonly represent our degree of belief that two taxa interact in nature, but can also represent the
182 probability of *observing* this interaction (Catchen *et al.* (2023)). For example, Gravel *et al.* (2019) used a
183 dataset of binary local European food webs of willow-galling sawflies and their natural enemies, all referenced
184 in space and time and consisting of similar species, to infer the probabilities of locally observing interactions
185 between co-occurring species. This was achieved by situating local webs within the context of environmental
186 variables, i.e. by including temperature and precipitation as conditional variables in some of their models.

187 We define space as the collection of geographic coordinates (x, y, z) , with (x, y) representing longitude and
188 latitude coordinates, and z denoting either altitudes or depths. These point coordinates delineate the spatial
189 boundaries of the system, which may be portrayed as a polyhedron. Ecological interactions may exhibit
190 variations along altitudinal gradients, as evidenced by changes in hummingbird-plant interactions (Weinstein &
191 Graham (2017b) and Weinstein & Graham (2017a)) and mosquito biting rates (e.g., Kulkarni *et al.* (2006)) at
192 different elevations. In contrast, time is defined as the specific time period within which interactions were either
193 observed or predicted. With these definitions in mind, space and time can be conceptualized as distinct patches
194 or time segments. Treating them as discrete dimensions aligns with the common sampling methods of
195 ecological webs and provides actual probabilities of interactions across space and time. This differs from the
196 approach of treating them as continuous variables, where edge values represent probability densities (i.e.,
197 relative likelihoods of interactions occurring at infinitesimal locations and instants in time) rather than discrete
198 probability values (which can be obtained by integrating probability densities). By employing discrete locations
199 and time periods, we can quantify an area (or volume) A and a duration t , which can be readily used in
200 spatiotemporal analyses of ecological networks. For example, when studying network-area relationships (NAR,
201 Galiana *et al.* (2018)), we anticipate that local probabilities of interactions scale positively with area and
202 duration because taxa have more opportunities to interact as these dimensions expand.

203 The probability that two taxa i and j interact in a local web $L_{x,y,z,t}$ (spatial and temporal subscripts hereafter

204 omitted for clarity) can be conditional on many environmental and biological factors. One of these is their
205 co-occurrence $C_{i,j}$, which is usually a Boolean describing if the geographic distributions of both taxa overlap
206 within the study area. Co-occurrence can be modeled probabilistically, in which case it may conform to a
207 Bernoulli distribution $C_{i,j} \sim \text{Bernoulli}(P(C_{i,j}))$, where the probability of co-occurrence $P(C_{i,j})$ can be
208 estimated through the application of joint species distribution models (e.g., Pollock *et al.* (2014)). Considering
209 that the probability that two non-co-occurring taxa interact locally is zero (i.e., $P(L_{i \rightarrow j}|C_{i,j} = 0) = 0$), the
210 probability of local interaction can be obtained by multiplying the probability of interaction given
211 co-occurrence with the probability of co-occurrence:

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

212 Other important factors that can impact interaction probabilities at the local scale are taxa relative abundance
213 (Canard *et al.* (2012)) and traits (Poisot *et al.* (2015)), as well as environmental factors such as temperature
214 (Angilletta *et al.* (2004)), precipitation (Woodward *et al.* (2012)), habitat structure (Klecka & Boukal (2014)),
215 and presence of other interacting taxa in the network (Pilosof *et al.* (2017), Kéfi *et al.* (2012)), as described
216 above. Here, we use the variable $\Omega_{x,y,z,t}$ (hereafter simply Ω) to describe the biological and ecological context
217 in which interaction probabilities were estimated. For example, if a research team conducts a mesocosm
218 experiment to estimate interaction probabilities between predators and prey with and without shelters (a place
219 that offers refuge and protection for prey, shielding them from predators), Ω would represent the presence or
220 absence of these shelters. Like co-occurrence, Ω can also be modeled probabilistically when the stochasticity or
221 uncertainty of environmental and biological factors is considered. In sum, Ω represents all ecological and
222 biological variables that were taken into consideration when measuring interaction probabilities and is,
223 therefore, a subset of all factors actually impacting ecological interactions.

224 The probability that two taxa i and j interact in a local web L can thus be conditional on the area (or volume) A ,
225 the time interval t , their co-occurrence $C_{i,j}$ and chosen environmental and biological conditions Ω . Although
226 these variables are associated with distinct questions or mechanisms related to ecological interactions, they may
227 covary with each other, such as the possible dependence of $C_{i,j}$ and Ω on spatial and temporal scales. When
228 estimating interaction probabilities using e.g. a generalized linear model with multiple explanatory variables
229 that might not be independent, it may become important to address collinearity. Using variable selection
230 techniques, for instance, may be necessary before fitting the model to mitigate this issue. The probability of

231 local interaction is described by the following expression when all these conditional variables are included:

$$P(L_{i \rightarrow j} | A, t, C_{i,j}, \Omega) \quad (2)$$

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

Table 1: Mathematical expression 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 L of realized interactions. Each expression 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.

Expression	Type	Outcome	Reference
$P(M_{i \rightarrow j})$	regional	biological feasibility of the interaction	Strydom <i>et al.</i> (2022)
$P(L_{i \rightarrow j})$	local	realization of the interaction	Fortuna & Bascompte (2006) (null model)
$P(L_{i \rightarrow j} A)$	local	realization of the interaction in a given area or volume	Galiana <i>et al.</i> (2018) *
$P(L_{i \rightarrow j} t)$	local	realization of the interaction during a given time period	Weinstein & Graham (2017a)

Expression	Type	Outcome	Reference
$P(L_{i \rightarrow j} C_{i,j})$	local	realization of the interaction given that the taxa co-occur	Gravel <i>et al.</i> (2019)
$P(L_{i \rightarrow j} \Omega)$	local	realization of the interaction given environmental conditions	Gravel <i>et al.</i> (2019) (temperature and precipitation)
$P(L_{i \rightarrow j} M_{i \rightarrow j})$	local	realization of the interaction given that the taxa can biologically interact	this study

244 Metawebs: regional catalogs of interactions

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

$$P(M_{i \rightarrow j}), \quad (3)$$

which, in contrast with local webs, is never conditional on any spatial, temporal, co-occurrence 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., $P(L_{i \rightarrow j}) < 1$) 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 webs

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

289 **Host-parasite network data**

290 We use the collection of tripartite host-parasite networks sampled across Europe, created by Kopelke *et al.*
291 (2017), in most of our case studies. This dataset contains well-resolved binary local interactions between
292 willows (52 species), willow-galling sawflies (96 species), and their parasitoids (126 species). Given its
293 replicated webs spanning large spatiotemporal scales, this dataset is well-suited for analyzing a variety of
294 ecological hypotheses and processes. Out of a total of 374 local webs, we retained those containing at least 5
295 species, resulting in a set of 233 georeferenced local webs (networks sampled within areas of 0.1 to 0.3 km²
296 during June and/or July spanning 29 years). We built a binary metaweb by aggregating all local interactions,
297 which gave us a regional web composed of 274 species and 1080 interactions. In the first two panels of fig. 1,
298 we show how the dissimilarity of interactions between common species (β_{OS}) and the dissimilarity in species
299 composition (β_S) between the metaweb and aggregated local webs (Poisot *et al.* (2012)) vary with the number
300 of sampled local webs. This shows that networks of local interactions are highly dissimilar from the metaweb,
301 both in terms of species and interactions, especially when only a limited number of sites has been sampled.
302 Both dissimilarity indices were calculated based on the number of items shared by the two webs (c_{LM}) and the
303 number of items unique to the metaweb (u_M) and to the aggregated local web (u_L). The β_S dissimilarity index
304 uses species (nodes) as items being compared, while the β_{OS} index assesses dissimilarity based on interactions
305 between shared species (Poisot *et al.* (2012)). Both indices were calculated following the β_W index of
306 Whittaker (1960):

$$\beta_W = \frac{c_{LM} + u_L + u_M}{(2c_{LM} + u_L + u_M)/2} - 1. \quad (4)$$

307 [Figure 1 about here.]

308 We converted these binary networks into probabilistic ones using models based on simple assumptions. Our
309 models are not designed to estimate the exact values of probabilistic interactions. Instead, their purpose is to
310 create plausible networks that serve as illustrative examples to highlight distinctions between probabilistic local
311 and metawebs. We created two probabilistic metawebs by employing constant false positive and false negative
312 rates for all regional interactions. In the first metaweb, we set both false positive and false negative rates to zero
313 to prevent artificially inflating the total number of links, enabling a more accurate comparison with binary webs.
314 This gave us a probability of regional interaction of 1 when at least one interaction has been observed and of 0

315 in the absence of any observed interaction between a given pair of species. In the second metaweb, we
 316 introduced a 5% false positive rate to account for spurious interactions and a 10% false negative rate to address
 317 the elevated occurrence of missing interactions in ecological networks (Catchen *et al.* (2023)). We believe these
 318 rates represent reasonable estimates of missing and spurious potential interactions, but confirming their
 319 accuracy is challenging due to the unavailability of data on the actual feasibility of interaction. Observed
 320 interactions were thus given a probability of regional interaction of 95%, whereas unobserved ones were
 321 assigned a probability of 10%.

322 To build probabilistic local webs, we first recognize that local interactions must initially be biologically feasible
 323 before occurring at a specific time and space. A local probability of interaction $P(L_{i \rightarrow j})$ can be expressed as the
 324 product of the probability of local interaction given that the two taxa can potentially interact $P(L_{i \rightarrow j}|M_{i \rightarrow j} = 1)$,
 325 which we denote as $P(L|M)$ for the sake of simplicity, with their probability of regional interaction $P(M_{i \rightarrow j})$:

$$P(L_{i \rightarrow j}) = P(L_{i \rightarrow j}|M_{i \rightarrow j} = 1) \times P(M_{i \rightarrow j} = 1). \quad (5)$$

326 We built the probabilistic local webs from the binary ones by using the probabilistic metawebs and a constant
 327 value of $P(L|M)$ across interactions. Low values of $P(L|M)$ indicate that feasible interactions rarely occur
 328 locally, intermediate values around 50% suggest considerable spatiotemporal variability, while high values
 329 indicate that regional interactions are nearly always realized locally. Following eq. 5, the local probability of
 330 interaction between a given pair of taxa consistently remained equal to or below their probability of regional
 331 interaction.

332 In the last two panels of fig. 1, we show how the aggregated number of links and connectance (i.e., the
 333 proportion of all of the non-forbidden links that are realized) scale with the number of sampled local
 334 probabilistic webs, according to different values of $P(L|M)$. When aggregating local probabilistic webs, the
 335 constancy of the probability of regional interaction across the entire study area means that any rise in the
 336 probability of local interaction is solely attributable to an increase in $P(L|M)$. The probability $P(L_{1,2}|M)$ of
 337 local interaction among potentially interacting species in an aggregated web $L_{1,2}$ is obtained by:

$$P(L_{1,2}|M) = 1 - (1 - P(L_1|M)) \times (1 - P(L_2|M)), \quad (6)$$

338 where $P(L_1|M)$ and $P(L_2|M)$ are the probabilities of local interaction among two potentially interacting species

339 in the subnetworks L_1 and L_2 , respectively. This equation provides the probability that the interaction is realized
340 in either (1) exclusively the local web L_1 , (2) exclusively the local web L_2 or (3) both local webs, given that the
341 two taxa have the biological capacity to interact.

342 By comparing the scaling relationships observed in binary and probabilistic webs, fig. 1 illustrates that high
343 values of $P(L|M)$ lead to systematic overestimations in the number of links and connectance, especially when
344 $P(L|M) = 1$ (corresponding to the scenario where local probabilities of interactions are equivalent to the
345 probabilities of regional interactions). However, these biases tend to diminish as the number of sampled webs
346 increases.

347 Taxonomic scaling of interactions

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

356 There are no inherent differences in the taxonomic scaling between local and metawebs, as only the nodes are
357 defined taxonomically. In other words, the probability values of edges in both local (eq. 2) and metawebs (eq. 3)
358 are not conditional on any taxonomic scale. The phylogenetic scale is tied to the definition of the event itself
359 (i.e., the interaction between two taxa), not to the conditional variables. In both types of webs, transitioning to a
360 broader level of organization (e.g., from a species-level web S to a genus-level web G) can be accomplished
361 directly by using probabilities from finer scales. For example, in a network with n_1 species from genus g_1 and
362 n_2 species from genus g_2 , one can compute the probability that at least one species from genus g_1 interacts with
363 at least one species from genus g_2 as follows:

$$P(G_{g_1 \rightarrow g_2}) = 1 - \prod_{i=1}^{n_1} \prod_{j=1}^{n_2} (1 - P(S_{g_1 i \rightarrow g_2 j})), \quad (7)$$

364 where $g_1 i$ and $g_2 j$ are the species of the corresponding genus. If it is known that at least two of these species

365 interact (i.e., $P(S_{g_{1i} \rightarrow g_{2j}}) = 1$ for at least one pair of (g_{1i}, g_{2j})), it implies a probability of genus interaction
366 equal to 1. Canard *et al.* (2012) built a species-based network following a similar approach, by using simulated
367 interactions between individuals derived from a neutral model (i.e., a model that assumed ecological
368 equivalence among individuals). In contrast, a more sophisticated approach is necessary when transitioning
369 from a broader to a finer level of organization. This is because knowledge of an interaction between two genera
370 does not guarantee that all possible pairwise combinations of their species will also interact. One possible
371 method is to build a finer-scale network by generating probabilities of interactions through random sampling
372 from a beta distribution, parameterized by the broader-scale network.

373 Ideally, our biological interpretation of probabilistic interactions should remain consistent across a network
374 even if it incorporates heterogeneous levels of organization, such as a network whose nodes represent both
375 species and trophic species (groups of species sharing similar predators and prey). This is common in
376 ecological webs where taxonomic resolution is typically low (Hemprich-Bennett *et al.* (2021), Vázquez *et al.*
377 (2022)). Interaction probabilities at broader taxonomic scales should be based on probabilities of interactions
378 between individuals, for both local and metawebs. For instance, in local individual-based food webs, the
379 probability that two individuals interact reflects our degree of belief that one individual will consume the other.
380 Similarly, in local species-based food webs, the probability that two species interact represents our degree of
381 belief that at least one individual from the predator species consumes at least another individual from the prey
382 species. Furthermore, in local clade-based food webs, the probability that two clades interact represents our
383 degree of belief that at least two species from these clades interact with each other or, equivalently, that at least
384 two individuals from these clades interact with each other. Fundamentally, the taxonomic scaling of interactions
385 involves aggregating interactions between individuals into larger groups, which may exhibit varying degrees of
386 homogeneity based on the organisms and the study system. In that regard, taxonomic scaling is analogous to the
387 spatial and temporal scaling of interactions, as they all represent different ways to aggregate individuals into
388 broader groups (either spatially, temporally, or taxonomically).

389 **Spatial and temporal scaling of interactions**

390 The investigation of network-area relationships and interaction accumulation curves is an important area of
391 research in network ecology. First, network-area relationships elucidate the scaling of network properties (such
392 as modularity and trophic chain lengths) with spatial scale (e.g., Wood *et al.* (2015), Galiana *et al.* (2018)). The
393 variations in network structure across spatial scales may stem from the scaling of species richness (species-area

relationships, SARs) and the number of interactions (Brose *et al.* (2004)) with the sampled area. Additionally, ecological processes unfolding at distinct spatial scales, such as the spatial variability in local community composition resulting from different sequences of extinction and colonization events, can also contribute to this variation (Galiana *et al.* (2018)). Next, interaction accumulation curves describe the scaling of the number of observed interactions with sampling effort (Jordano (2016)). Sampling effort, which may correspond to the duration of the sampling period used to construct the network, can impact connectance (Bersier *et al.* (1999)) and various measures of network structure (Banašek-Richter *et al.* (2004), McLeod *et al.* (2021)). Apart from sampling effort, the temporal scaling of interactions also elucidates how network structure changes with the temporal resolution of the network (Poisot *et al.* (2012)), acknowledging that distinct interactions take place over time, ranging from short-term fluctuations of interactions to long-term trends. As probabilistic local webs may explicitly account for the spatiotemporal variability of interactions, they offer a distinct approach to investigating the scaling of network structure with space and time, in contrast to binary and quantitative webs, by making the stochasticity of interactions the focal point of the modeling process.

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

On the other hand, local interactions scale both spatially and temporally, given that they have more opportunities to be realized and observed in larger areas and longer durations. This is attributed to factors such as a higher number of individuals, greater trait variations, and increased opportunities for encounters, as highlighted by McLeod *et al.* (2020). For example, if a local probabilistic web L_1 with an area A_1 is derived

424 from a larger web L_0 with an area A_0 , and A_1 is entirely nested within A_0 , interaction probabilities should be
425 lower in the smaller web, i.e. $P(L_{1,i \rightarrow j}|A_1 < A_0) \leq P(L_{0,i \rightarrow j}|A_0)$. However, if A_1 and A_0 are disjoint, interaction
426 probabilities could be higher in the smaller area, contingent upon their environmental and biological conditions.
427 Likewise, interaction probabilities are expected to be lower in webs with shorter durations when time intervals
428 are nested. In fig. 2, we show how the expected number of local host-parasite interactions scales with area,
429 represented as an expanding latitudinal window, in comparison with regional interactions. Even though we
430 employed local probabilities of interactions equal to regional interactions for the purpose of comparison (i.e.,
431 using $P(L|M) = 1$ here), we notice that the total number of regional interactions scales more rapidly than local
432 interactions. This is because numerous regional interactions involve species that never co-occur, and as a result,
433 these interactions are not captured in local webs.

434 [Figure 2 about here.]

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

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

451 As indicated by eq. 1, the probability that two taxa i and j interact at a specific location (x, y, z) is determined by

452 the product of their probability of interaction given co-occurrence and their probability of co-occurrence. First,
 453 their co-occurrence probability can be calculated using their individual occurrence probabilities $P(X_i)$ and
 454 $P(X_j)$, along with the strength of association γ between their occurrences and co-occurrence (Cazelles *et al.*
 455 (2016)):

$$P(C_{i,j}) = P(X_i, X_j) = P(X_i)P(X_j)\gamma. \quad (8)$$

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

$$C_{i,j} \sim \text{Bernoulli}(P(X_i, X_j)). \quad (9)$$

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

$$P(L_{i \rightarrow j}|C_{i,j} = 1) = 1 - e^{-\lambda t_0}, \quad (10)$$

475 which tends toward 1 as $t_0 \rightarrow \infty$. It is important to note that the units of λ and t_0 are complementary. For
 476 instance, if the duration t_0 is measured in months, λ denote the expected number of interactions per month.
 477 The occurrence of an interaction between i and j can be modeled as a Bernoulli trial with a probability of
 478 $P(L_{i \rightarrow j})$. Consequently, a Bayesian inference model can be built based on the preceding equations to estimate
 479 the value of the λ and γ parameters and generate novel interaction data:

$$L_{i \rightarrow j} \sim \text{Bernoulli}(P(L_{i \rightarrow j})) \quad (11)$$

$$P(L_{i \rightarrow j}) = P(X_i)P(X_j)\gamma(1 - e^{-\lambda t_0}) \quad (12)$$

$$\gamma \sim \text{Gamma}(2, 0.5) \quad (13)$$

$$\lambda \sim \text{Exponential}(2) \quad (14)$$

470 In fig. 3, we show the variation in the probability of interaction under different parameter values. In the right
 471 panel, we notice that, irrespective of the interaction rate λ , the probability of interaction converges toward an
 472 asymptote determined by the co-occurrence $P(C_{i,j})$ (eq. 8). This model can be customized in different ways,
 473 such as by linking λ with specific environmental variables or explicitly incorporating observation errors (i.e.,
 474 the probabilities of false negatives and false positives).

475 [Figure 3 about here.]

476 Sampling binary webs

477 The prediction of binary interactions through random draws is an important application of probabilistic webs.
 478 This approach proves beneficial for analyzing the structural characteristics of probabilistic webs, particularly in
 479 the absence of specific analytical measures. By simulating independent Bernoulli trials for each interaction, a
 480 binary network can be generated. A probability distribution of network properties can be obtained by measuring
 481 network structure across multiple randomly generated networks (Poisot *et al.* (2016)). This method enables the
 482 representation of the variability of network structure, albeit with possible biases when connectance is low
 483 (Poisot & Gravel (2014), Chagnon (2015)). Employing this strategy to generate binary networks under a null
 484 model facilitates null hypothesis significance testing, wherein the observed measure is compared against the
 485 simulated distribution (e.g., Bascompte *et al.* (2003)). Additionally, randomly generating binary networks
 486 across spatial and temporal dimensions aids in representing the spatiotemporal variability of network structure,
 487 allowing the testing of ecological hypotheses regarding interactions on large spatial and temporal scales.
 488 There are at least two distinct approaches to sample binary networks from probabilistic webs across space, for
 489 example, when attempting to predict a binary network for each of a number of locations within a given region.

490 The first approach involves performing a singular Bernoulli trial for each pair of taxa based on their regional
 491 probability of interaction. In employing this approach, every pair of taxa predicted to interact in the binary
 492 metaweb will be treated as interacting in all local webs where they co-occur. This will result in local pairwise
 493 interactions without spatial variation. The second approach is to independently sample each of the local
 494 probabilistic webs, which can be achieved by first generating distinct probabilistic networks for each location.
 495 These local probabilistic webs may vary in taxa composition and interaction probabilities. Subsequently, binary
 496 networks can be independently sampled for each location, introducing spatial variation in binary interactions.
 497 In fig. 4, we compare the average connectance of binary networks resulting from these two sampling techniques,
 498 where regional and local interactions are drawn from our host-parasite probabilistic networks, generating a
 499 number of binary web realizations for each site in the dataset. These two sampling approaches yield different
 500 outcomes, particularly for lower values of $P(L|M)$, which denote instances where regional interactions do not
 501 consistently manifest locally. Small discrepancies are also apparent between these techniques when we equate
 502 the probability of local interaction to the probability of regional interaction (i.e., when using $P(L|M) = 1.0$ in
 503 eq. 5), especially when the number of binary network samples for each location is low. As anticipated, we
 504 observe that sampling binary interactions from the metaweb tends to overestimate connectance on average
 505 compared to sampling them from local webs. Furthermore, we observe an increase in the variability of
 506 connectance when employing a single sample, representing what we consider as a more tangible process
 507 leading to the realization of local and regional interactions in nature.

508 [Figure 4 about here.]

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

$$P(L_{0,i \rightarrow j}) = 1 - (1 - P(L_{1,i \rightarrow j}) \times P(L_{2,i \rightarrow j})). \quad (15)$$

515 Due to its larger area, the probability that the two taxa interact in L_0 is equal or greater than in L_1 and L_2 . When

516 sampling binary interactions from local webs, it is crucial to sample at the same spatial scale for which
517 probabilities were estimated. Otherwise, interaction probabilities must be adjusted to align with the intended
518 cell size, preventing systematic biases in predictions.

519 **Prediction of local webs from metawebs**

520 Metawebs serve as a valuable source of ecological information for predicting local webs across time and space.
521 Binary local webs can be reconstructed by selecting a subset of taxa and interactions from the metaweb (Dunne
522 (2006)). This implies that metawebs consistently contain more interactions than their corresponding local webs.
523 Determining the list of taxa to select can be achieved empirically or through methods like range maps or species
524 distribution models. As the species composition of a community is arguably less difficult to sample or predict
525 than its interactions, the primary challenge lies in deciding which interactions to select from the metaweb, a task
526 that may necessitate advanced statistical models and ecological expertise. Inferring the structure of local webs
527 from the metaweb before predicting specific local pairwise interactions could hold promise (Strydom *et al.* (2021)),
528 considering that the structure of local webs is constrained by the metaweb (Saravia *et al.* (2022)).
529 Building probabilistic local webs from a probabilistic metaweb involves a reduction in the value of pairwise
530 interaction probabilities. This decrease is attributed to the prerequisite that two taxa must initially possess the
531 capacity to interact before engaging in local interactions (eq. 5). Therefore, inferring local webs from their
532 metaweb while maintaining identical interaction probability values would introduce systematic biases into the
533 predictions. In such cases, these networks would essentially represent smaller-scale metawebs of potential
534 interactions, possibly leading to misinterpretations by being perceived as local interactions. As proposed by
535 McLeod *et al.* (2021), although metawebs do not capture the spatiotemporal variability of interactions, they
536 establish an upper limit for local interactions. In other words, the probability of two taxa interacting at a specific
537 location and time is consistently lower or equal to the probability of their regional interaction:

$$P(L_{i \rightarrow j} | A, t, C, \Omega) \leq P(M_{i \rightarrow j}). \quad (16)$$

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

$$\int_{\Omega} \int_A \int_t P(L_{i \rightarrow j} | A, t, \Omega) dt dA d\Omega \leq P(M_{i \rightarrow j}). \quad (17)$$

542 Estimating more precisely the probability $P(L|M)$ that two taxa interact locally if they can potentially interact
 543 allows for improved predictions of local webs from a probabilistic metaweb. This task is challenging due to the
 544 variability of this probability across space and time, as well as its variability across pairwise interactions within
 545 a network. Using simple models of $P(L|M)$, as demonstrated in our case studies, represents an initial step
 546 toward the overarching objective of reconstructing local webs from metawebs.

547 Conclusion

548 In this contribution, we underline the importance of network metadata for adequately interpreting and
 549 manipulating probabilistic interaction data. The mathematical representation of probabilities and their statistical
 550 properties depend on the type of interactions (local or regional) and the conditions under which these
 551 interactions were evaluated. We showed that probabilistic local and metawebs differ in their relationship to
 552 spatial and temporal scales, with regional interactions remaining consistent across scales. In contrast with
 553 metawebs, local interactions are measured in a specific context (e.g., in a given area, time, and biological and
 554 environmental conditions) and depend on taxa co-occurrence. These important conceptual differences bring to
 555 light the need to use probabilistic data with caution, for instance when generating binary network realizations
 556 across space and predicting local webs from metawebs. Clear metadata describing the type of interaction and
 557 the variables used in their estimation are required to ensure adequate data manipulation. Better data practices
 558 and foundations for probabilistic thinking in network ecology could enable more reliable assessments of the
 559 spatiotemporal variability and uncertainty of biotic interactions.
 560 It is essential to enhance our comprehension of both regional and local interactions, especially considering the
 561 current scarcity of interaction data. However, while sampling biological communities does decrease the
 562 uncertainty of interactions by accumulating evidence for their feasibility and local realization, there is a limit to
 563 how much we can diminish their inherent randomness. In metawebs, probabilities reflect our limited knowledge
 564 of interactions (i.e., our degree of belief that interactions are feasible), which is expected to improve with a
 565 larger volume of data. We should anticipate that regional interactions will become more definitive (with
 566 probabilities approaching 0 or 1) as we investigate various conditions, including different combinations of

567 species traits. Conversely, in the case of local webs, which can be seen as random instances of metawebs,
568 randomness cannot be as reduced. In local interactions, probabilities can be divided into a component
569 representing uncertainty and another representing spatiotemporal variability. Owing to environmental
570 heterogeneity, there will invariably be instances in which an interaction occurs and others in which it does not,
571 across different times and locations, irrespective of the extent to which we can improve our knowledge of its
572 biological feasibility and the local conditions that facilitate its occurrence. When local webs depict probabilities
573 of observing interactions rather than just their actual occurrence, we must also consider the observation
574 variability as an additional source of randomness. Every ecological process is stochastic but there is also a
575 possibility that a phenomenon goes undetected. Quantifying and partitioning this stochasticity will enable us to
576 make more accurate predictions about ecological interactions at various spatial and temporal scales. This will
577 prove to be of vital importance as our time to understand nature runs out, especially at the places where the
578 impacts of climate change and habitat loss hit harder.

579 Acknowledgment

580 We acknowledge that this study was conducted on land within the traditional unceded territory of the Saint
581 Lawrence Iroquoian, Anishinabewaki, Mohawk, Huron-Wendat, and Omàmiwininiwak nations. This work was
582 supported by the Institute for Data Valorisation (IVADO) and the Natural Sciences and Engineering Research
583 Council of Canada (NSERC) Collaborative Research and Training Experience (CREATE) program, through the
584 Computational Biodiversity Science and Services (BIOS²) program. A special thanks to all members of the
585 Black Holes and Revelations working group (organized by BIOS²) for their insightful discussions and valuable
586 feedback on this manuscript.

587 References

- 588 Abrego, N., Roslin, T., Huotari, T., Ji, Y., Schmidt, N.M., Wang, J., *et al.* (2021). Accounting for species
589 interactions is necessary for predicting how arctic arthropod communities respond to climate change.
590 *Ecography*, 44, 885–896.
- 591 Andrade-Pacheco, R., Rerolle, F., Lemoine, J., Hernandez, L., Meïté, A., Juziwelo, L., *et al.* (2020). Finding
592 hotspots: Development of an adaptive spatial sampling approach. *Scientific Reports*, 10, 10939.

- 593 Angilletta, M.J., Jr., Steury, T.D. & Sears, M.W. (2004). Temperature, Growth Rate, and Body Size in
594 Ectotherms: Fitting Pieces of a Life-History Puzzle¹. *Integrative and Comparative Biology*, 44, 498–509.
- 595 Araujo, M.B., Rozenfeld, A., Rahbek, C. & Marquet, P.A. (2011). Using species co-occurrence networks to
596 assess the impacts of climate change. *Ecography*, 34, 897–908.
- 597 Banašek-Richter, C., Cattin, M.-F. & Bersier, L.-F. (2004). Sampling effects and the robustness of quantitative
598 and qualitative food-web descriptors. *Journal of Theoretical Biology*, 226, 23–32.
- 599 Bartomeus, I., Gravel, D., Tylianakis, J.M., Aizen, M.A., Dickie, I.A. & Bernard-Verdier, M. (2016). A
600 common framework for identifying linkage rules across different types of interactions. *Functional Ecology*,
601 30, 1894–1903.
- 602 Bascompte, J., Jordano, P., Melián, C.J. & Olesen, J.M. (2003). The nested assembly of plantanimal mutualistic
603 networks. *Proceedings of the National Academy of Sciences*, 100, 9383–9387.
- 604 Berlow, E.L., Neutel, A.-M., Cohen, J.E., De Ruiter, P.C., Ebenman, B., Emmerson, M., et al. (2004).
605 Interaction strengths in food webs: Issues and opportunities. *Journal of Animal Ecology*, 73, 585–598.
- 606 Bersier, L.-F., Dixon, P. & Sugihara, G. (1999). Scale-Invariant or Scale-Dependent Behavior of the Link
607 Density Property in Food Webs: A Matter of Sampling Effort? *The American Naturalist*, 153, 676–682.
- 608 Blanchet, F.G., Cazelles, K. & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions.
609 *Ecology Letters*, 23, 1050–1063.
- 610 Bohan, D.A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A.J. & Woodward, G. (2017).
611 Next-Generation Global Biomonitoring: Large-scale, Automated Reconstruction of Ecological Networks.
612 *Trends in Ecology & Evolution*, 32, 477–487.
- 613 Bolnick, D.I., Amarasekare, P., Araújo, M.S., Bürger, R., Levine, J.M., Novak, M., et al. (2011). Why
614 intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution*, 26, 183–192.
- 615 Borrett, S.R. & Scharler, U.M. (2019). Walk partitions of flow in Ecological Network Analysis: Review and
616 synthesis of methods and indicators. *Ecological Indicators*, 106, 105451.
- 617 Boulangeat, I., Gravel, D. & Thuiller, W. (2012). Accounting for dispersal and biotic interactions to disentangle
618 the drivers of species distributions and their abundances. *Ecology Letters*, 15, 584–593.
- 619 Brimacombe, C., Bodner, K., Michalska-Smith, M., Poisot, T. & Fortin, M.-J. (2023). Shortcomings of reusing
620 species interaction networks created by different sets of researchers. *PLOS Biology*, 21, e3002068.

- 621 Broom, M. & Ruxton, G.D. (2005). You can run or you can hide: Optimal strategies for cryptic prey against
622 pursuit predators. *Behavioral Ecology*, 16, 534–540.
- 623 Brose, U., Ostling, A., Harrison, K. & Martinez, N.D. (2004). Unified spatial scaling of species and their
624 trophic interactions. *Nature*, 428, 167–171.
- 625 Canard, E., Mouquet, N., Marescot, L., Gaston, K.J., Gravel, D. & Mouillot, D. (2012). Emergence of
626 Structural Patterns in Neutral Trophic Networks. *PLOS ONE*, 7, e38295.
- 627 Catchen, M.D., Poisot, T., Pollock, L.J. & Gonzalez, A. (2023). The missing link: Discerning true from false
628 negatives when sampling species interaction networks.
- 629 Cazelles, K., Araujo, M.B., Mouquet, N. & Gravel, D. (2016). A theory for species co-occurrence in interaction
630 networks. *Theoretical Ecology*, 9, 39–48.
- 631 Chagnon, P.-L. (2015). Characterizing topology of ecological networks along gradients: The limits of metrics'
632 standardization. *Ecological Complexity*, 22, 36–39.
- 633 Choh, Y., Ignacio, M., Sabelis, M.W. & Janssen, A. (2012). Predator-prey role reversals, juvenile experience
634 and adult antipredator behaviour. *Scientific Reports*, 2, 728.
- 635 Cirtwill, A.R., Eklöf, A., Roslin, T., Wootton, K. & Gravel, D. (2019). A quantitative framework for
636 investigating the reliability of empirical network construction. *Methods in Ecology and Evolution*, 10,
637 902–911.
- 638 Dallas, T., Park, A.W. & Drake, J.M. (2017). Predicting cryptic links in host-parasite networks. *PLOS
639 Computational Biology*, 13, e1005557.
- 640 Dunne, J.A. (2006). The Network Structure of Food Webs. In: *Ecological networks: Linking structure and
641 dynamics* (eds. Dunne, J.A. & Pascual, M.). Oxford University Press, pp. 27–86.
- 642 Elmasri, M., Farrell, M.J., Davies, T.J. & Stephens, D.A. (2020). A hierarchical Bayesian model for predicting
643 ecological interactions using scaled evolutionary relationships. *The Annals of Applied Statistics*, 14,
644 221–240.
- 645 Elton, C.S. (2001). *Animal Ecology*. University of Chicago Press, Chicago, IL.
- 646 Emmerson, M.C. & Raffaelli, D. (2004). Predator-prey body size, interaction strength and the stability of a real
647 food web. *Journal of Animal Ecology*, 73, 399–409.

- 648 Fortuna, M.A. & Bascompte, J. (2006). [Habitat loss and the structure of plantanimal mutualistic networks](#).
649 *Ecology Letters*, 9, 281–286.
- 650 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*,
651 33, 3117–3128.
- 653 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.
- 655 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.
- 657 Godsoe, W., Murray, R. & Iritani, R. (2022). [Species interactions and diversity: A unified framework using Hill numbers](#). *Oikos*, n/a, e09282.
- 659 Golubski, A.J. & Abrams, P.A. (2011). [Modifying modifiers: What happens when interspecific interactions interact?](#) *Journal of Animal Ecology*, 80, 1097–1108.
- 661 Gomez, J.M., Verdu, M. & Perfectti, F. (2010). [Ecological interactions are evolutionarily conserved across the entire tree of life](#). *Nature*, 465, 918–U6.
- 663 Gonzalez, A. & Londoño, M.C. (2022). [Monitor biodiversity for action](#). *Science*, 378, 1147–1147.
- 664 Gonzalez-Varo, J.P. & Traveset, A. (2016). [The Labile Limits of Forbidden Interactions](#). *Trends in Ecology & Evolution*, 31, 700–710.
- 666 Gravel, D., Baiser, B., Dunne, J.A., Kopelke, J.-P., Martinez, N.D., Nyman, T., *et al.* (2019). [Bringing Elton and Grinnell together: A quantitative framework to represent the biogeography of ecological interaction networks](#). *Ecography*, 42, 401–415.
- 669 Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. (2013). [Inferring food web structure from predatorprey body size relationships](#). *Methods in Ecology and Evolution*, 4, 1083–1090.
- 671 Guimarães, P.R. (2020). [The Structure of Ecological Networks Across Levels of Organization](#). *Annual Review of Ecology, Evolution, and Systematics*, 51, 433–460.
- 673 Guimerà, R. & Sales-Pardo, M. (2009). [Missing and spurious interactions and the reconstruction of complex networks](#). *Proceedings of the National Academy of Sciences*, 106, 22073–22078.

- 675 Hemprich-Bennett, D.R., Oliveira, H.F.M., Le Comber, S.C., Rossiter, S.J. & Clare, E.L. (2021). [Assessing the](#)
676 [impact of taxon resolution on network structure](#). *Ecology*, 102, e03256.
- 677 Herrera, C.M. (1989). [Pollinator abundance, morphology, and flower visitation rate: Analysis of the “quantity”](#)
678 [component in a plant-pollinator system](#). *Oecologia*, 80, 241–248.
- 679 Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M. & Ladle, R.J. (2015). [Seven Shortfalls](#)
680 [that Beset Large-Scale Knowledge of Biodiversity](#). *Annual Review of Ecology, Evolution, and Systematics*,
681 46, 523–549.
- 682 Jordano, P. (1987). [Patterns of Mutualistic Interactions in Pollination and Seed Dispersal: Connectance,](#)
683 [Dependence Asymmetries, and Coevolution](#). *The American Naturalist*, 129, 657–677.
- 684 Jordano, P. (2016). [Sampling networks of ecological interactions](#). *Functional Ecology*, 30, 1883–1893.
- 685 Jordano, P., Bascompte, J. & Olesen, J.M. (2003). [Invariant properties in coevolutionary networks of](#)
686 [plantanimal interactions](#). *Ecology Letters*, 6, 69–81.
- 687 Kéfi, S., Berlow, E.L., Wieters, E.A., Joppa, L.N., Wood, S.A., Brose, U., *et al.* (2015). [Network structure](#)
688 [beyond food webs: Mapping non-trophic and trophic interactions on Chilean rocky shores](#). *Ecology*, 96,
689 291–303.
- 690 Kéfi, S., Berlow, E.L., Wieters, E.A., Navarrete, S.A., Petchey, O.L., Wood, S.A., *et al.* (2012). [More than a](#)
691 [meal... integrating non-feeding interactions into food webs](#). *Ecology Letters*, 15, 291–300.
- 692 Kéfi, S., Miele, V., Wieters, E.A., Navarrete, S.A. & Berlow, E.L. (2016). [How Structured Is the Entangled](#)
693 [Bank? The Surprisingly Simple Organization of Multiplex Ecological Networks Leads to Increased](#)
694 [Persistence and Resilience](#). *PLOS Biology*, 14, e1002527.
- 695 Klecka, J. & Boukal, D.S. (2014). [The effect of habitat structure on prey mortality depends on predator and](#)
696 [prey microhabitat use](#). *Oecologia*, 176, 183–191.
- 697 Kopelke, J.-P., Nyman, T., Cazelles, K., Gravel, D., Vissault, S. & Roslin, T. (2017). [Food-web structure of](#)
698 [willow-galling sawflies and their natural enemies across Europe](#). *Ecology*, 98, 1730–1730.
- 699 Kulkarni, M.A., Kweka, E., Nyale, E., Lyatuu, E., Mosha, F.W., Chandramohan, D., *et al.* (2006).
700 [Entomological Evaluation of Malaria Vectors at Different Altitudes in Hai District, Northeastern Tanzania](#).
701 *Journal of Medical Entomology*, 43, 580–588.
- 702 McCann, K. (2007). Protecting biostructure. *Nature*, 446, 29–29.

- 703 McCann, K.S. (2011). Food webs (MPB-50). In: *Food Webs (MPB-50)*. Princeton University Press.
- 704 McLeod, A., Leroux, S.J., Gravel, D., Chu, C., Cirtwill, A.R., Fortin, M.-J., et al. (2021). [Sampling and](#)
705 [asymptotic network properties of spatial multi-trophic networks](#). *Oikos*, n/a.
- 706 McLeod, A.M., Leroux, S.J. & Chu, C. (2020). [Effects of species traits, motif profiles, and environment on](#)
707 [spatial variation in multi-trophic antagonistic networks](#). *Ecosphere*, 11, e03018.
- 708 Melián, C.J., Vilas, C., Baldó, F., González-Ortegón, E., Drake, P. & Williams, R.J. (2011). [Eco-evolutionary](#)
709 [Dynamics of Individual-Based Food Webs](#). In: *Advances in Ecological Research*, The Role of Body Size in
710 Multispecies Systems (ed. Belgrano, A.). Academic Press, pp. 225–268.
- 711 Ohlmann, M., Miele, V., Dray, S., Chalmandrier, L., O'Connor, L. & Thuiller, W. (2019). [Diversity indices for](#)
712 [ecological networks: A unifying framework using Hill numbers](#). *Ecology Letters*, 22, 737–747.
- 713 Olesen, J.M., Bascompte, J., Dupont, Y.L., Elberling, H., Rasmussen, C. & Jordano, P. (2010). [Missing and](#)
714 [forbidden links in mutualistic networks](#). *Proceedings of the Royal Society B: Biological Sciences*, 278,
715 725–732.
- 716 Pascual, M., Dunne, J.A. & Dunne, J.A. (2006). *Ecological Networks: Linking Structure to Dynamics in Food*
717 *Webs*. Oxford University Press, USA.
- 718 Petchey, O.L., Beckerman, A.P., Riede, J.O. & Warren, P.H. (2008). [Size, foraging, and food web structure](#).
719 *Proceedings of the National Academy of Sciences*, 105, 4191–4196.
- 720 Pilosof, S., Porter, M.A., Pascual, M. & Kéfi, S. (2017). [The multilayer nature of ecological networks](#). *Nature*
721 *Ecology & Evolution*, 1, 1–9.
- 722 Plagányi, É.E. & Butterworth, D.S. (2004). [A critical look at the potential of Ecopath with ecosim to assist in](#)
723 [practical fisheries management](#). *African Journal of Marine Science*, 26, 261–287.
- 724 Poisot, T., Canard, E., Mouillot, D., Mouquet, N. & Gravel, D. (2012). [The dissimilarity of species interaction](#)
725 [networks](#). *Ecology Letters*, 15, 1353–1361.
- 726 Poisot, T., Cirtwill, A.R., Cazelles, K., Gravel, D., Fortin, M.-J. & Stouffer, D.B. (2016). [The structure of](#)
727 [probabilistic networks](#). *Methods in Ecology and Evolution*, 7, 303–312.
- 728 Poisot, T. & Gravel, D. (2014). [When is an ecological network complex? Connectance drives degree](#)
729 [distribution and emerging network properties](#). *PeerJ*, 2, e251.

- 730 Poisot, T., Stouffer, D.B. & Gravel, D. (2015). [Beyond species: Why ecological interaction networks vary](#)
731 [through space and time](#). *Oikos*, 124, 243–251.
- 732 Polis, G.A. (1991). [Complex Trophic Interactions in Deserts: An Empirical Critique of Food-Web Theory](#). *The*
733 [American Naturalist](#), 138, 123–155.
- 734 Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O’Hara, R.B., Parris, K.M., *et al.* (2014). [Understanding](#)
735 [co-occurrence bymodelling species simultaneously with a Joint Species DistributionModel \(JSDM\)](#).
736 [Methods in Ecology and Evolution](#), 5, 397–406.
- 737 Pomeranz, J.P.F., Thompson, R.M., Poisot, T. & Harding, J.S. (2019). [Inferring predatorprey interactions in](#)
738 [food webs](#). *Methods in Ecology and Evolution*, 10, 356–367.
- 739 Pringle, R.M. & Hutchinson, M.C. (2020). [Resolving Food-Web Structure](#). *Annual Review of Ecology,*
740 [Evolution, and Systematics](#), 51, 55–80.
- 741 Proulx, S.R., Promislow, D.E.L. & Phillips, P.C. (2005). [Network thinking in ecology and evolution](#). *Trends in*
742 [Ecology & Evolution](#), SPECIAL ISSUE: BUMPER BOOK REVIEW, 20, 345–353.
- 743 Pulliam, H.R. (1974). [On the Theory of Optimal Diets](#). *The American Naturalist*, 108, 59–74.
- 744 Rooney, N. & McCann, K.S. (2012). [Integrating food web diversity, structure and stability](#). *Trends in Ecology*
745 [& Evolution](#), 27, 40–46.
- 746 Salim, J.A., Saraiva, A.M., Zermoglio, P.F., Agostini, K., Wolowski, M., Drucker, D.P., *et al.* (2022). [Data](#)
747 [standardization of plantpollinator interactions](#). *GigaScience*, 11, giac043.
- 748 Sanders, D. & van Veen, F.J.F. (2012). [Indirect commensalism promotes persistence of secondary consumer](#)
749 [species](#). *Biology Letters*, 8, 960–963.
- 750 Saravia, L.A., Marina, T.I., Kristensen, N.P., De Troch, M. & Momo, F.R. (2022). [Ecological network](#)
751 [assembly: How the regional metaweb influences local food webs](#). *Journal of Animal Ecology*, 91, 630–642.
- 752 Schmeller, D.S., Julliard, R., Bellingham, P.J., Böhm, M., Brummitt, N., Chiarucci, A., *et al.* (2015). [Towards a](#)
753 [global terrestrial species monitoring program](#). *Journal for Nature Conservation*, 25, 51–57.
- 754 Singer, M.C. & McBride, C.S. (2012). [Geographic mosaics of species’ association: A definition and an](#)
755 [example driven by plantinsect phenological synchrony](#). *Ecology*, 93, 2658–2673.

- 756 Stouffer, D.B., Rezende, E.L. & Amaral, L.A.N. (2011). [The role of body mass in diet contiguity and food-web](#)
757 [structure](#). *Journal of Animal Ecology*, 80, 632–639.
- 758 Strydom, T., Bouskila, S., Banville, F., Barros, C., Caron, D., Farrell, M.J., *et al.* (2022). [Food web](#)
759 [reconstruction through phylogenetic transfer of low-rank network representation](#). *Methods in Ecology and*
760 *Evolution*, 13.
- 761 Strydom, T., Catchen, M.D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., *et al.* (2021). [A](#)
762 [roadmap towards predicting species interaction networks \(across space and time\)](#). *Philosophical*
763 *Transactions of the Royal Society B-Biological Sciences*, 376, 20210063.
- 764 Tylianakis, J.M., Tscharntke, T. & Lewis, O.T. (2007). [Habitat modification alters the structure of tropical](#)
765 [hostparasitoid food webs](#). *Nature*, 445, 202–205.
- 766 Vázquez, D.P., Peralta, G., Cagnolo, L., Santos, M. & Igual, E. autores contribuyeron por. (2022). [Ecological](#)
767 [interaction networks. What we know, what we don't, and why it matters](#). *Ecología Austral*, 32, 670–697.
- 768 Weinstein, B.G. & Graham, C.H. (2017a). [On comparing traits and abundance for predicting species](#)
769 [interactions with imperfect detection](#). *Food Webs*, 11, 17–25.
- 770 Weinstein, B.G. & Graham, C.H. (2017b). [Persistent bill and corolla matching despite shifting temporal](#)
771 [resources in tropical hummingbird-plant interactions](#). *Ecology Letters*, 20, 326–335.
- 772 Whittaker, R.H. (1960). [Vegetation of the Siskiyou Mountains, Oregon and California](#). *Ecological*
773 *Monographs*, 30, 279–338.
- 774 Wood, S.A., Russell, R., Hanson, D., Williams, R.J. & Dunne, J.A. (2015). [Effects of spatial scale of sampling](#)
775 [on food web structure](#). *Ecology and Evolution*, 5, 3769–3782.
- 776 Woodward, G., Benstead, J.P., Beveridge, O.S., Blanchard, J., Brey, T., Brown, L.E., *et al.* (2010). [Chapter 2 -](#)
777 [Ecological Networks in a Changing Climate](#). In: *Advances in Ecological Research*, Ecological Networks
778 (ed. Woodward, G.). Academic Press, pp. 71–138.
- 779 Woodward, G., Brown, L.E., Edwards, F.K., Hudson, L.N., Milner, A.M., Reuman, D.C., *et al.* (2012). [Climate](#)
780 [change impacts in multispecies systems: Drought alters food web size structure in a field experiment](#).
- 781 [Philosophical Transactions of the Royal Society B: Biological Sciences, 367, 2990–2997.](#)

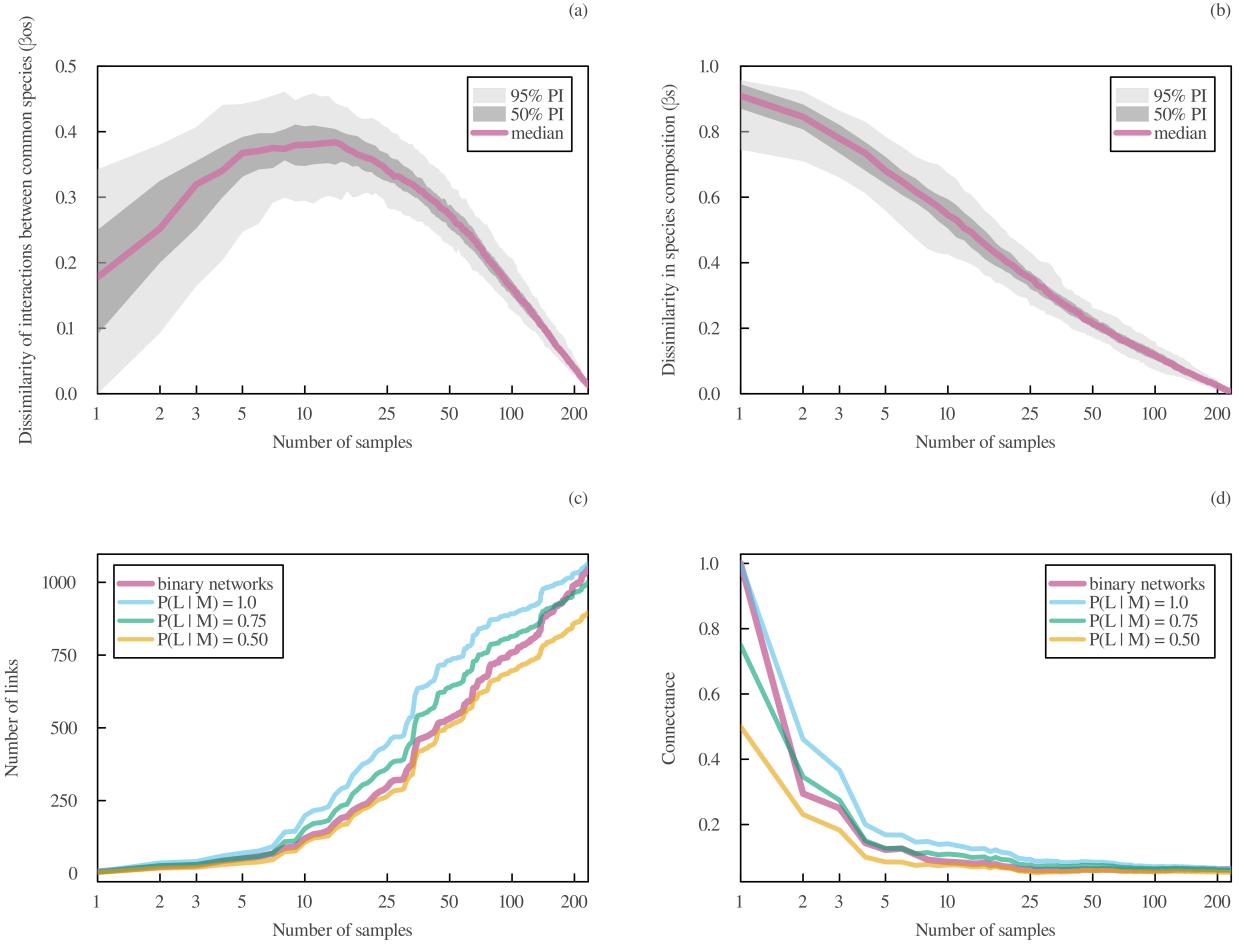


Figure 1: Network accumulation curves. (a) Dissimilarity of interactions between common species and (b) dissimilarity in species composition between aggregated local webs and the binary metaweb of host-parasite interactions. Aggregated local webs were obtained by sequentially and randomly selecting a 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 $P(L|M)$ (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 $P(L|M)$ adjusting according to eq. 6).

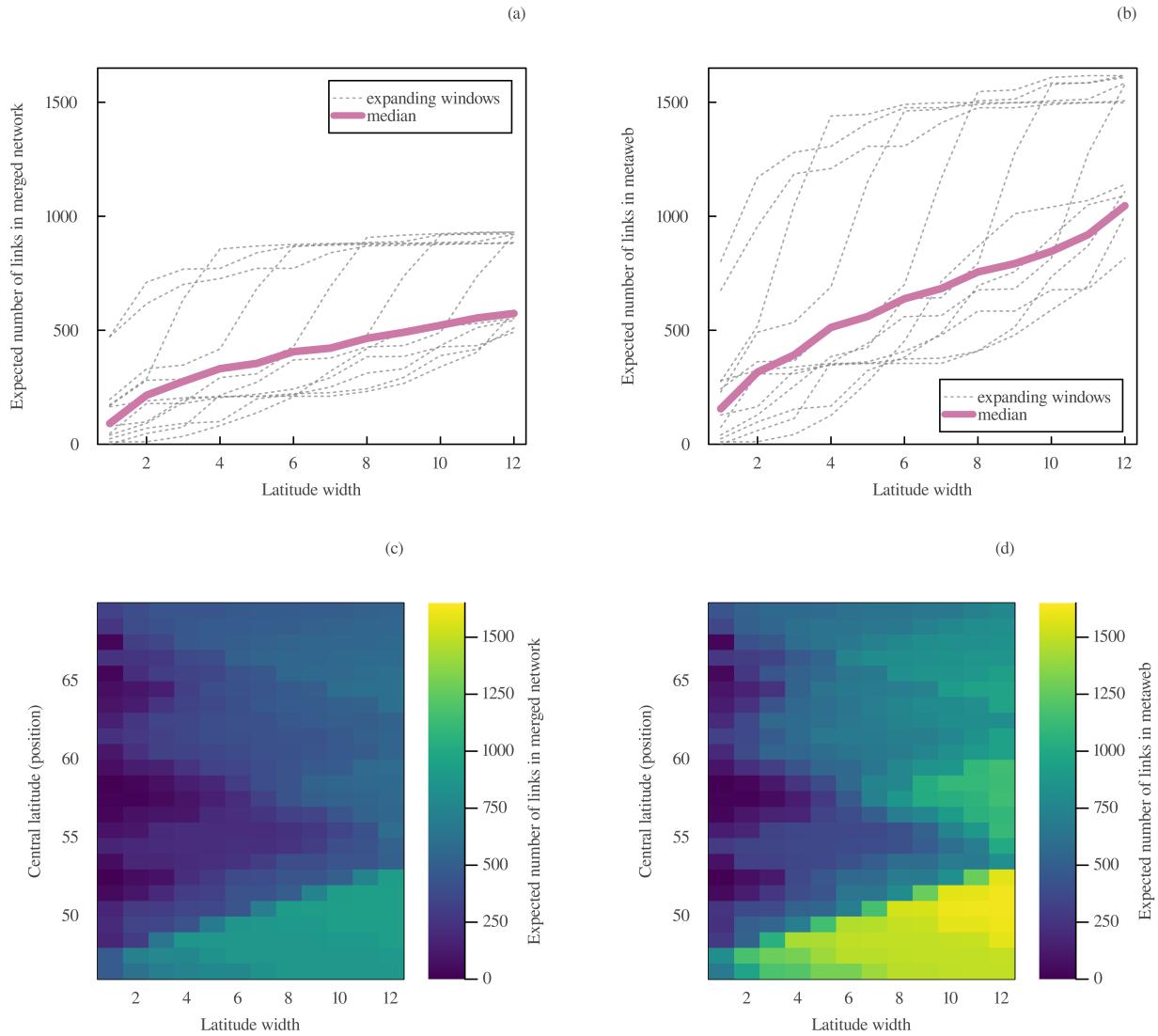


Figure 2: Spatial scaling of interactions. Expected number of host-parasite interactions in a network aggregating all probabilistic (a) local and (b) regional interactions within a latitudinal window of a given length. Every dashed curve corresponds to a different expanding window characterized by a 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 $P(L|M)$ (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 $P(L|M)$ remaining at their maximum value of 1 following eq. 6.

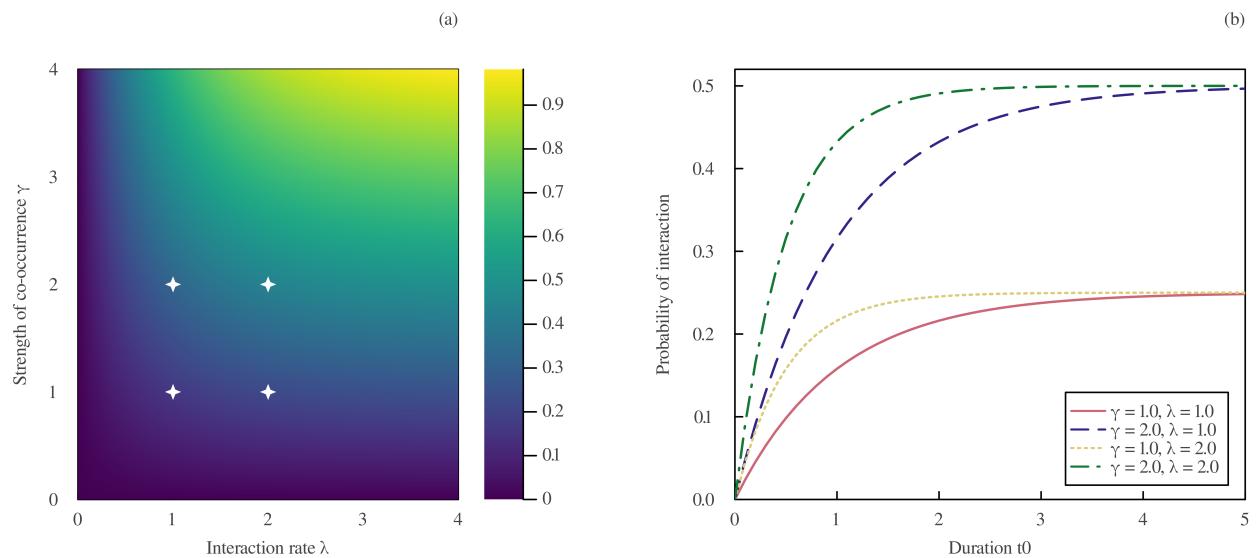


Figure 3: **Parameters of the spatiotemporally explicit model of interactions.** (a) Probability of local interaction given by the process model (eq. 12) under different values of λ and γ , with $t_0 = 1$. The parameter values used in the right panel are denoted by the white stars. (b) Scaling of the probability of interaction with the duration parameter t_0 in eq. 12, for different values of λ and γ . 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.

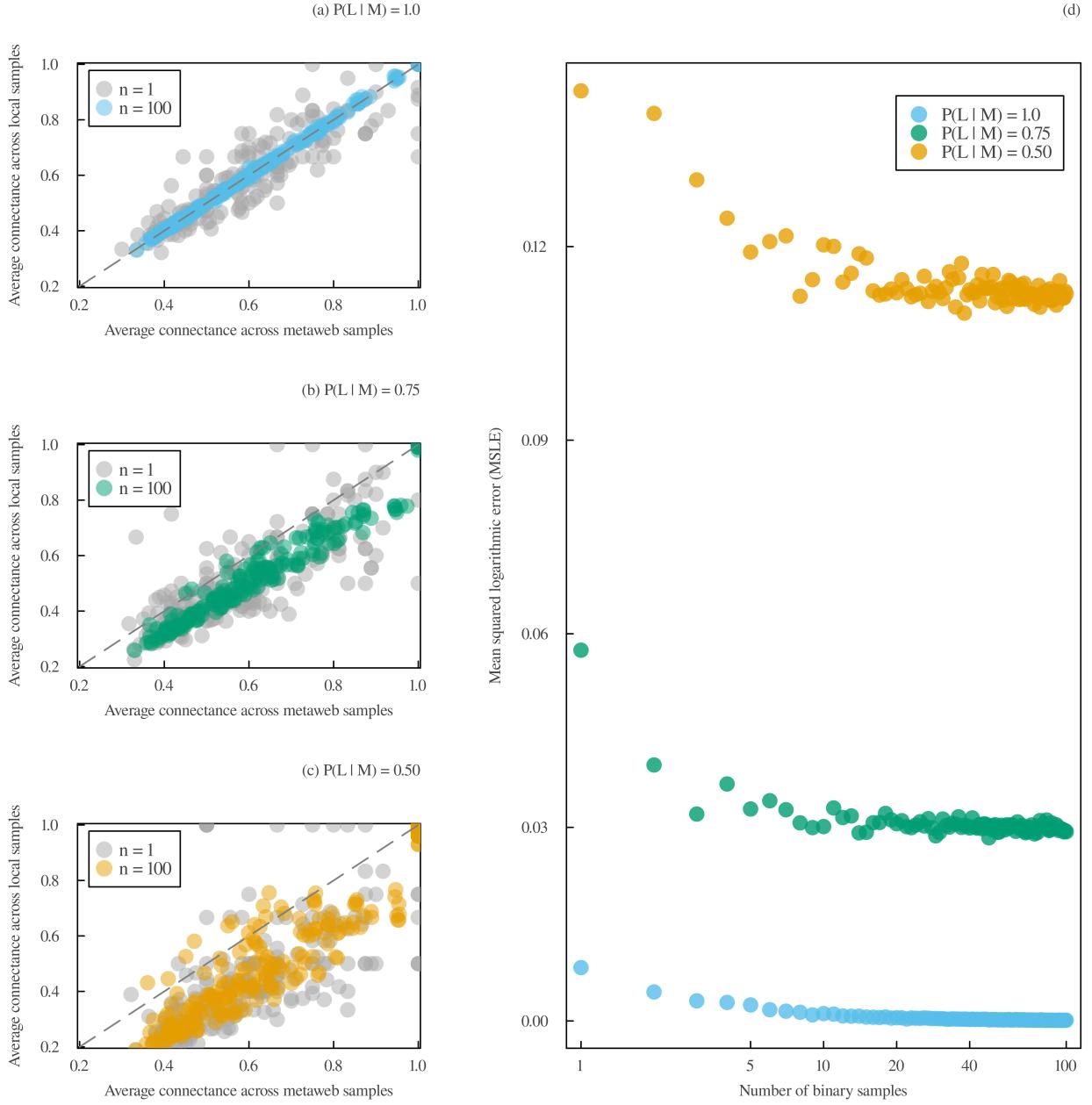


Figure 4: Connectance of sampled binary webs. Comparison between the average connectance of binary network samples obtained from the probabilistic local and metawebs. Each dot corresponds to a different site. The local probability of interaction between potentially interacting species was set to three different values: (a) $P(L|M) = 1.0$, (b) $P(L|M) = 0.75$, and (c) $P(L|M) = 0.50$. 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 $P(L|M)$ 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.