

Deciphering ecological networks of probabilistic interactions

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 are they to occur?) as opposed to deterministically (are they occurring?) highlights uncertainties in our knowledge of interactions. The origin of the uncertainty captured by interaction probabilities differs depending on the method used to evaluate them. It may consist of the uncertainty of predictive models, the subjective assessment of experts, or even the empirical evaluation of interaction spatiotemporal variability. Despite this growing interest in probabilistic interactions, guidelines for the estimation and documentation of probabilistic interaction data are still lacking. This is concerning because our biological understanding of interaction probabilities and the decisions made during data analysis rely heavily on the often elusive definition of the probabilistic event. 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 (predatory, herbivory, and parasitic interactions). These definitions are based on the distinction between the realization of an interaction at a specific time and space (local) and its biological feasibility (regional). Through case studies using open data on host-parasite interactions in Europe, we show that these two network representations have different statistical properties when it comes to common ecological applications. Specifically, we found that local and metawebs differ in their spatial and temporal scaling of probabilistic interactions, but not in their taxonomic scaling. We suggest two approaches to inferring binary interactions from probabilistic ones that account for these differences and show that systematic biases arise when directly inferring local webs from subsets of metawebs. Our results underscore the importance of clear metadata for networks of probabilistic interactions that describe the type of interaction (local or regional)

and uncertainty sources. Comprehensive metadata are particularly useful for the identification of optimal sampling locations where data collection may provide the greatest amount of new information on interactions and minimize their uncertainty.

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



Last revision: *March 5, 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 (e.g., nocturnal and diurnal species coexisting in the same space without interacting due to differences
9 in the timing of their daily activities, Jordano (1987)) and the large number of rare and cryptic interactions in a
10 community contribute to these uncertainties (Jordano (2016)). More generally, a handful of conditions must be
11 satisfied for an interaction to be observed locally. First, both species must have overlapping geographic ranges,
12 i.e. they must co-occur within the region of interest (Blanchet *et al.* (2020)). Second, they must have some
13 probability of meeting. Probabilities of interspecific encounters are typically low, especially for rare species
14 with low relative abundances (Canard *et al.* (2012)). The probability that species meet each other also depends
15 on their biological characteristics, such as the synchronization of their phenology (Olesen *et al.* (2010), Singer
16 & McBride (2012)) and their discoverability (e.g., Broom & Ruxton (2005)). Finally, when species do come
17 into contact, an interaction occurs only if their traits are locally compatible (Poisot *et al.* (2015)), including but
18 not limited to their body phenotypes (Bolnick *et al.* (2011), Stouffer *et al.* (2011), Gravel *et al.* (2013)) and
19 behavioral choices (Pulliam (1974), Choh *et al.* (2012)). Interactions may also be influenced by the presence or
20 prevalence of a third species (e.g., of a more profitable prey species, Golubski & Abrams (2011), Sanders & van
21 Veen (2012)). Documenting the location and timing of interactions becomes even more difficult when
22 accounting for the spatiotemporal variability of ecological interactions (Poisot *et al.* (2012), Poisot *et al.*
23 (2015)). Environmental factors, such as temperature (Angilletta *et al.* (2004)), drought (Woodward *et al.*
24 (2012)), climate change (Gilman *et al.* (2010), Woodward *et al.* (2010), Araujo *et al.* (2011)), and habitat
25 modifications (Tylianakis *et al.* (2007)), contribute to this spatiotemporal variability by impacting species
26 abundance and traits. Even after satisfying all these conditions, there remains a possibility that the interaction
27 does not occur locally, either due to the intricate nature of the system or simply by chance. If it does occur, it
28 might still go unnoticed, particularly if it happens infrequently. In this context, it is unsurprising that our
29 knowledge of ecological interactions remains limited (Hortal *et al.* (2015)) despite extensive biodiversity data
30 collection (Schmeller *et al.* (2015)).

31 Knowing the biological capacity of two species to interact directly (via e.g., trophic interactions) is necessary
32 but not sufficient for inferring their interaction at a specific time and space. The recognition of the intrinsic
33 variability of species interactions has led ecologists to expand their representation of ecological networks (also
34 known as ecological webs) to include a probabilistic view of interactions (Poisot *et al.* (2016), Dallas *et al.*
35 (2017), Fu *et al.* (2021)). This different perspective allows us to fill in the Eltonian shortfall (Hortal *et al.*
36 (2015)) by modeling the probability of detecting interactions, which can be an important tool for directing
37 efforts and taking action, especially in places where access and resources for research are scarce. The
38 probabilistic representation of interactions is thus far limited to direct interactions, which are conceptually and
39 mathematically analogous regardless of their biological type (e.g., predation and pollination). This is in contrast
40 with indirect interactions (e.g., interspecific competition), which arise from distinct ecological processes and are
41 often not directly observable (Kéfi *et al.* (2015), Kéfi *et al.* (2016)). Representing direct interactions
42 probabilistically enables us to capture the spatiotemporal variability of the aforementioned ecological processes
43 and the uncertainties associated with their measurement. As opposed to webs of binary deterministic
44 interactions, in which interactions are regarded as either occurring or not, webs of probabilistic interactions,
45 within a Bayesian framework, express our degree of belief (or confidence) regarding the occurrence or
46 observation of interactions. Based on the scale at which they are estimated, probabilistic interactions may
47 reflect our level of confidence in whether interactions will be observed, realized locally, or biologically feasible.
48 As an illustration, we could outline a situation in which there is a 50% certainty that an interaction occurs 50%
49 of the time, or that there is a 50% certainty that it simply occurs. Our level of confidence should be more
50 definitive (approaching either 0 or 1) as we extend our sampling to a broader area and over a longer time period,
51 thereby diminishing the uncertainty of the interactions (but not necessarily the estimation of their variability).
52 In the broadest sense, binary interactions are also a type of probabilistic interaction, in which the numerical
53 value of an interaction is restrained to 0 (non-occurring) or 1 (occurring). Yet, for the sake of clarity, we omit
54 binary interactions from our discussion of probabilistic interactions in this contribution. In networks of
55 probabilistic interactions, only forbidden interactions (i.e., interactions prohibited by biological traits or species
56 absence, Jordano *et al.* (2003), Olesen *et al.* (2010)) have a probability value of zero by default, provided that
57 intraspecific trait variability is considered (Gonzalez-Varo & Traveset (2016)).
58 By accounting for the uncertainty of interactions, networks of probabilistic interactions may provide a more
59 realistic portrait of species interactions and network structure (i.e. community-level properties), which are major
60 drivers of the functioning, dynamics, and resilience of ecosystems worldwide (Proulx *et al.* (2005), McCann

61 (2007), McCann (2011), Rooney & McCann (2012)). Moreover, the application and development of
62 computational methods in network ecology, often based on a probabilistic representation of interactions, can
63 alleviate (and guide) the sampling efforts required to document species interactions (Strydom *et al.* (2021)). For
64 example, statistical models can be used to estimate the uncertainty of pairwise interactions (Cirtwill *et al.*
65 (2019)) and the probability of missing (false negatives) and spurious (false positives) interactions (Guimerà &
66 Sales-Pardo (2009)). Considering the high rate of false negatives in species interaction data due to the difficulty
67 of witnessing rare interactions (Catchen *et al.* (2023)), these models could inform the identification of priority
68 sampling locations of ecological webs where data collection would yield the most valuable information, thereby
69 reducing errors. Optimization models for sampling locations have mostly found applications in biological
70 systems that are not networks, such as identifying priority sampling sites for disease hotspots (Andrade-Pacheco
71 *et al.* (2020)), but there is substantial promise in applying them to probabilistic ecological interactions.
72 Statistical models can also be used to generate predictions of ecological webs without prior knowledge of
73 pairwise interactions, for instance using body size (Petchey *et al.* (2008), Gravel *et al.* (2013)), phylogeny
74 (Elmasri *et al.* (2020), Strydom *et al.* (2022)), or a combination of niche and neutral processes (Bartomeus *et*
75 *al.* (2016), Pomeranz *et al.* (2019)) for inference. Topological null models, which generate networks of
76 probabilistic interactions by preserving chosen characteristics of the adjacency matrix of binary interactions
77 while intentionally omitting others (Bascompte *et al.* (2003), Fortuna & Bascompte (2006)), serve as other
78 examples of common probabilistic interaction models. Null models can be used to produce underlying
79 distributions of network measures for null hypothesis significance testing. Many measures have been developed
80 to describe the structure (Poisot *et al.* (2016)) and diversity (Ohlmann *et al.* (2019), Godsoe *et al.* (2022)) of
81 probabilistic interaction webs. These models and measures support the use of this approach for the study of a
82 wide range of ecological questions, from making better predictions of species distribution (Cazelles *et al.*
83 (2016)) to forecasting the impact of climate change on ecological webs (Gilman *et al.* (2010)).
84 The lack of clear guidelines on the use of probabilistic interaction data is worrisome, both for data producers
85 and re-users who generate and manipulate these numbers. This is concerning because sampling strategies and
86 decisions regarding network construction can affect our understanding of network properties (Brimacombe *et*
87 *al.* (2023)). Besides methodological difficulties that may arise when assessing probabilistic interactions, a
88 precise definition of probabilistic interactions appears to be lacking, making the estimation and use of these data
89 more difficult. We aim to take a step back by outlining different ways in which probabilistic interactions are
90 defined and used in network ecology. We distinguish two broad categories of probabilistic interaction networks

91 that necessitate distinct approaches when applied to key ecological questions: local webs describing
92 probabilities of realized interactions, and regional webs (metawebs) describing probabilities of potential
93 interactions. We highlight the distinctions in the ecological meaning of these two representations and examine
94 some of their properties and relationships (particularly with space, time, and between each other) through
95 empirical case studies. Moreover, there is currently no metadata standard that could guide the documentation of
96 all types of probabilistic interactions (although see e.g., Salim *et al.* (2022) who discuss data standards for
97 deterministic mutualistic webs). Well-defined metadata for probabilistic interactions would support more
98 adequate manipulation and integration of interaction data from different sources and guard against possible
99 misinterpretations arising from ambiguous definitions of probabilistic interaction networks. These metadata
100 should outline the nature (i.e., local or regional) and type (e.g., predatory or pollination) of the interactions,
101 provide information regarding the taxonomic level, identities, and characteristics (e.g., life stages) of the
102 individuals involved in an interaction, present the mathematical formulation of probabilities, including clearly
103 identified conditional variables (e.g., spatial and temporal scales), and describe the methods and contexts (e.g.,
104 location, time, environmental conditions) in which interactions were estimated. Inadequately documented
105 probabilistic interaction data should be used with caution when analyzing ecological webs. Our observations
106 and advice can be applied to most types of ecological networks representing direct interactions, from food webs
107 to host-virus networks.

108 **Probabilistic representations of interactions**

109 One of the first aspects to take into consideration when estimating or interpreting probabilities of interactions is
110 knowing if they describe potential or realized interactions, as these two types of interactions have distinct
111 conceptual underpinnings and sources of uncertainty. A potential (regional) interaction is defined as the
112 biological capacity of two taxa to interact (i.e., the probability that they *can* theoretically interact) whereas a
113 realized (local) interaction is the materialization or observation of this interaction in a well-defined space and
114 time (i.e., the probability that they interact locally). Our discussion of local interactions focus on their
115 occurrence rather than their direct empirical observation, as one of the primary goals of describing them
116 probabilistically is to characterize our uncertainty regarding their actual manifestation. We use the terms
117 *metaweb* (Dunne (2006)) to designate regional webs of potential interactions and *local webs* (Poisot *et al.*
118 (2012)) for those of realized interactions. Metawebs are the network analogs of the species pool, where local

119 webs originate from a subset of both species (nodes) and interactions (edges) of the regional metaweb (Saravia
120 *et al.* (2022)). Without clear documentation, it can be challenging to know if published probabilistic interaction
121 webs describe local or regional interactions (tbl. 1 provides examples of studies employing both types of
122 probabilistic interaction networks), or if so-called probabilities are in reality *interaction scores* (i.e.,
123 non-probabilistic quantitative interactions analogous to interaction strengths). When probabilistic regional
124 interactions are used and interpreted incorrectly as local interactions (and conversely), this may generate
125 misleading findings during data analysis. We believe that a better understanding of the differences, similarities,
126 and interconnections between these two probabilistic representations of ecological interactions would alleviate
127 interpretation errors (e.g., when studying network-area relationships) and facilitate a more adequate utilization
128 of interaction data.

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

130 Local and metawebs, like any type of network, are made of nodes and edges that can be represented at different
131 levels of organization and precision. The basic units of ecological networks are individuals that interact with
132 each other (e.g., by predation in food webs, Elton (2001)), forming individual-based networks (Melián *et al.*
133 (2011)). The aggregation of these individuals into more or less homogeneous groups (e.g., populations, species,
134 families, feeding guilds) allows us to represent nodes at broader taxonomic scales, which affects our
135 interpretation of the properties of these systems (Guimarães (2020), Hemprich-Bennett *et al.* (2021)).

136 Edges linking nodes can describe a variety of interaction measures. Ecologists have traditionally represented
137 interactions as binary objects that were considered realized after observing at least one individual from group i
138 interact with at least another individual from group j . In an adjacency matrix of binary interactions B , the
139 presence or absence of an interaction $B_{i \rightarrow j}$ between two taxa can be viewed as the result of a Bernoulli process
140 $B_{i \rightarrow j} \sim \text{Bernoulli}(P(B_{i \rightarrow j}))$, with $P(B_{i \rightarrow j})$ being the probability of interaction that characterizes our limited
141 knowledge of the system and its intrinsic spatiotemporal variability. In networks of probabilistic interactions,
142 $P(B_{i \rightarrow j})$ are edge values, and the only two possible outcomes are the presence ($B_{i \rightarrow j} = 1$) or absence ($B_{i \rightarrow j} = 0$)
143 of an interaction between each pair of nodes. Depending on the type of probabilistic interaction network (local
144 or metaweb), the mathematical formulation and interpretation of stochastic parameters like $P(B_{i \rightarrow j})$ can be
145 linked to environmental and biological factors such as species relative abundance, traits, area, and time (tbl. 1),
146 for example using logistic regression with a logit link function with continuous explanatory variables. Predicting
147 the number of local webs in which the interaction occurs can be achieved by using a Binomial distribution,

148 assuming a constant probability of interaction and independence between networks (trials). When considering
149 uncertainties around the estimation of $P(B_{i \rightarrow j})$, a Beta distribution can also be used to encompass all possible
150 probability values. In that case, a Beta-Binomial distribution can be used to predict the number of networks in
151 which the interaction occurs. Observing an interaction between two taxa at a given location and time provides
152 important information that can be used to update previous estimates of $P(B_{i \rightarrow j})$, informing us on the biological
153 capacity of both taxa to interact and the environmental conditions that enabled them to interact locally.

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

166 Just like binary interaction networks, the uncertainty and spatiotemporal variability of interaction strengths can
167 be represented probabilistically. However, the need to estimate the probability distribution of all possible values
168 of interaction strengths can make the inference of probabilities more challenging in quantitative webs compared
169 to binary interaction webs, which require only one probability estimate for each interaction. Interaction
170 strengths can follow various probability distributions depending on the measure used. For instance, they can
171 follow a Poisson distribution $W_{i \rightarrow j} \sim \text{Poisson}(\lambda_{i \rightarrow j})$ when predicting frequencies of interactions between pairs
172 of nodes, with $\lambda_{i \rightarrow j}$ being the expected rate at which individuals of taxa i and j interact (e.g., the average
173 number of prey j consumed by all predators i in a given time period). The Poisson distribution can also be
174 0-inflated after initially modeling non-interacting taxa (e.g., Boulangeat *et al.* (2012) employ a 0-inflated model
175 to analyze species abundance following the modeling of species presence and absence), which constitute the
176 majority of taxa pairs in most local webs due to their typically high sparseness (Jordano (2016)). Because of the
177 methodological difficulties typically encountered when building deterministic quantitative webs (which are only

178 partially mitigated by models such as Ecopath, Plagányi & Butterworth (2004)), binary interaction webs, which
179 are easier to sample (Jordano (2016)) and predict (Strydom *et al.* (2021)), have been more frequently studied
180 and modeled. Moreover, most published networks of probabilistic interactions and methods describe
181 probabilistic interactions whose outcome is binary (whether interaction probabilities are regarded as constant or
182 variable, e.g. represented by a Beta distribution), which underlines the need for better guidelines regarding the
183 interpretation and manipulation of these types of webs first. For these reasons, our primary focus in this
184 contribution will be on addressing the challenges in interpreting and using interaction probabilities in Bernoulli
185 distributions, in both local and metawebs.

186 Local webs: communities interacting in space and time

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

196 We define space as the collection of geographic coordinates (x, y, z) , with (x, y) representing longitude and
197 latitude coordinates, and z denoting either altitudes or depths. These point coordinates delineate the spatial
198 boundaries of the system, which may be portrayed as a polyhedron. Ecological interactions may exhibit
199 variations along latitudinal and altitudinal gradients, as evidenced by changes in hummingbird-plant
200 interactions (Weinstein & Graham (2017b), Weinstein & Graham (2017a)) and mosquito biting rates (e.g.,
201 Kulkarni *et al.* (2006)) at different elevations. In contrast, time is defined as the specific time period within
202 which interactions were either observed or predicted. Even though space and time are continuous variables that
203 should yield probability *densities* of interactions (i.e., relative likelihoods of interactions occurring at
204 infinitesimal locations and instants in time), these definitions enable them to be conceptualised as distinct
205 patches and time segments. Treating space and time as discrete dimensions aligns with the common sampling
206 methods of ecological networks and provides actual probabilities of interactions, which can be obtained by

integrating probability densities. Furthermore, we can quantify both an area A and a duration t , which can be readily used in spatiotemporal analyses of ecological networks. For example, when studying network-area relationships (NAR, Galiana *et al.* (2018)), we anticipate that local probabilities of interactions scale positively with area and duration because taxa have more opportunities to interact as these dimensions expand.

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

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

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

The probability that two taxa i and j interact in a local web L can thus be conditional on the area (or volume) A ,

234 the time interval t , their co-occurrence $C_{i,j}$ and chosen environmental and biological conditions Ω . Although
235 these variables are associated with distinct questions or mechanisms related to ecological interactions, they may
236 covary with each other, such as the possible dependence of $C_{i,j}$ and Ω on spatial and temporal scales. When
237 estimating interaction probabilities using e.g. a generalized linear model with multiple explanatory variables
238 that might not be independent, it may become important to address collinearity. Using variable selection
239 techniques, for instance, may be necessary before fitting the model to mitigate this issue. The probability of
240 local interaction is described by the following expression when all these conditional variables are included:

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

241 The representation of the local context in which probabilities are estimated and the variables that should be
242 taken into consideration depend on the study system, the objectives of the study, and the resources available to
243 the researchers. In other words, these variables do not systematically need to be accounted for. For example, in
244 Gravel *et al.* (2019), for the purpose of model comparison, some models incorporated environmental variables
245 as conditional factors to estimate interaction probabilities, while others did not. When accounted for, these
246 variables should be clearly described in the documentation of the data (Brimacombe *et al.* (2023)),
247 preferentially in mathematical terms to avoid any confusion in their interpretation and to limit manipulation
248 errors during their re-use. For instance, ecologists should be explicit about their consideration of co-occurrence
249 in their estimation of local interaction probabilities. Indeed, it is important to specify if probability values are
250 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
251 interpretation and analysis of the data. In [tbl. 1](#), we present examples of studies that used these diverse
252 formulations of probabilistic interactions and conditional variables.

Table 1: Mathematical expressions of probabilistic interactions. The probability of interaction between two taxa i and j is interpreted differently in a metaweb M of potential interactions (representing the *biological* feasibility of interactions), a metaweb M^* of potential interactions (representing the *ecological* feasibility of interactions) and a local web L of realized interactions. Each expression emphasizes a different conditional variable, the three dots serving as a placeholder for other variables not explicitly stated in the expression. The outcome of each of these probabilistic events, along with common uncertainty sources (often stemming from the model used for estimation), is presented alongside examples of studies that employed them (with specific variables indicated in parentheses, when applicable). The study marked with an asterisk has been conducted on binary interaction webs. Note that local interaction probabilities may represent the probability of observing an interaction, not necessarily its actual occurrence.

Expression	Type	Outcome	Uncertainty sources	Reference
$P(M_{i \rightarrow j} T_i, T_j)$	regional	biological feasibility of the interaction given regional traits (non-forbiddenness)	trait matching models	Strydom <i>et al.</i> (2022)
$P(M_{i \rightarrow j}^* T_i, T_j, E)$	regional	ecological feasibility of the interaction given regional traits and environmental conditions	trait matching and environmental-based models	this study
$P(L_{i \rightarrow j} X_{il}, X_{jl}, \text{local})$		realization of the interaction given taxa co-occurrence	spatiotemporal variability	Gravel <i>et al.</i> (2019)
$P(L_{i \rightarrow j} N_{il}, N_{jl}, \text{local})$		realization of the interaction given taxa abundances	neutral models	Canard <i>et al.</i> (2014)
$P(L_{i \rightarrow j} T_{il}, T_{jl}, \text{local})$		realization of the interaction given local traits	trait matching models	Gravel <i>et al.</i> (2016)
$P(L_{i \rightarrow j} E_l, \dots)$	local	realization of the interaction given local environmental conditions	environmental-based models	Gravel <i>et al.</i> (2019) (temperature and precipitation)
$P(L_{i \rightarrow j} A, \dots)$	local	realization of the interaction in a given area or volume	spatial models	Galiana <i>et al.</i> (2018) *

Expression	Type	Outcome	Uncertainty sources	Reference
$P(L_{i \rightarrow j} t, \dots)$	local	realization of the interaction during a given time period	temporal models	Weinstein & Graham (2017a)
$P(L_{i \rightarrow j} f(L), \dots)$	local	realization of the interaction given network structure	common null models	Fortuna & Bascompte (2006) (connectance)
$P(L_{i \rightarrow j} M_{i \rightarrow j}, \dots)$	local	realization of the interaction given that the taxa can biologically interact	spatiotemporal variability	this study

263 Metawebs: regional catalogs of interactions

264 Metawebs are networks of potential interactions that have been designed for broad spatial, temporal, and
 265 taxonomic scales (e.g., species food webs at the continental scale). They represent the probability that taxa can
 266 biologically interact regardless of their co-occurrence and local environmental conditions. Metawebs of
 267 probabilistic interactions are particularly valuable when interaction data is limited, i.e. when the uncertainty of
 268 potential interactions is high. As data accumulates, interactions in metawebs should tend towards binarity,
 269 approaching probability values of 0 (repeatedly failing to observe an interaction) and 1 (observing an
 270 interaction at least once).

271 Regional interactions are by definition context-independent, i.e. they are not measured at a specific location and
 272 time. In contrast with local webs of probabilistic interactions, which represent the stochasticity of interactions
 273 occurring in nature, metawebs of probabilistic interactions measure our degree of belief in the capacity of two
 274 taxa to interact (i.e., the probability that their traits could support an interaction in the right conditions).
 275 Consequently, although *neutrally* forbidden interactions (i.e., improbable interactions between rare species,
 276 Canard *et al.* (2012)) tend to have low probability values in local webs, they may exhibit a higher probability in
 277 the metaweb. Potential interactions describe the probability that there exists at least one combination of
 278 phenotypes of taxa i and j that can interact with each other if they were to encounter. For example, a piscivorous
 279 species may have the capacity to interact with a potential prey species if they share at least one matching

270 combination of body sizes, such as an adult predator being able to consume juvenile prey. To reduce any biases
271 in our calculation of this probability, it is crucial to ensure that the set of traits sampled or considered accurately
272 reflects the overall trait distribution in both taxa. This enhances our confidence in concluding that the taxa
273 cannot interact when no positive observation has been made and increases our capacity to detect an interaction
274 when it is biologically feasible. In addition to pairwise interactions, the extent of sampling effort also
275 contributes to shaping our assessment of metaweb properties, as sampling a larger number of local webs allows
276 us to capture more regional interactions (McLeod *et al.* (2021)). A probability of potential interaction in a
277 metaweb M can be expressed as

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

278 which, in contrast with local webs, is never conditional on any spatial, temporal, co-occurrence or
279 environmental variables (tbl. 1).

280 Starting from a selected set of taxa, which are usually distributed within a broad region of interest, metawebs
281 can be built using different data sources, including literature review (e.g., Maiorano2020Tetraeuia), aggregated
282 local web data (e.g., Gravel *et al.* (2019), Saravia *et al.* (2022)), and predictive models (e.g., Strydom *et al.*
283 (2022), Shaw *et al.* (2024)). Every pair of taxa that have confidently been observed to interact at least once can
284 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
285 webs of probabilistic interactions, in which interaction events usually remain stochastic (i.e., $P(L_{i \rightarrow j}) < 1$) after
286 empirically observing interactions because of their intrinsic spatiotemporal variability. Similarly, although rare
287 interactions typically have low probabilities in local webs, they can have high probabilities in metawebs when
288 the traits of both taxa are congruent. On the other hand, interactions that were never observed can have low
289 probability values in both local and metawebs, going as low as 0 for forbidden links (i.e., interactions forbidden
290 by biological traits). However, many observations of interactions are in reality false positives because of
291 observation errors due to taxonomic misidentifications and ecological misinterpretations (e.g., due to cryptic
292 species and interactions, Pringle & Hutchinson (2020)). Likewise, forbidden interactions can be false negatives
293 in metawebs, e.g. if they have been assessed for specific phenotypes, locations or time. Employing Bayesian
294 models, whether they are mechanistic or phenomenological, could improve the accuracy of our estimation of
295 interaction probabilities within both local and regional webs (e.g., Bartomeus *et al.* (2016), Cirtwill *et al.*
296 (2019)). This improvement is achieved in metawebs by leveraging prior information regarding the feasibility of

297 interactions along with the empirical data on observed interactions.

298 Properties of probabilistic interaction webs

299 Local and metawebs of probabilistic interactions differ in their type of interactions (i.e., realized or potential)
300 and in the conditional variables upon which interaction values depend. These differences are significant as they
301 influence the characteristics of probabilistic interaction networks. Neglecting to consider them may result in
302 misleading results and interpretation errors when analyzing the properties of probabilistic interaction webs,
303 which could be particularly problematic when addressing crucial ecological questions about networks. Here we
304 compare the characteristics of local and metawebs through the presentation of four applications of probabilistic
305 interactions: (1) describing their spatial and temporal scaling, (2) describing their taxonomic scaling, (3)
306 sampling for binary interaction webs, and (4) reconstructing local webs of probabilistic interactions from
307 metawebs. All code and data to reproduce these analyses are available at the Open Science Framework (TBD).

308 Host-parasite network data

309 We use the collection of tripartite host-parasite networks sampled across Europe, created by Kopelke *et al.*
310 (2017), in the following case studies. This dataset contains well-resolved binary local interactions between
311 willows (52 species), willow-galling sawflies (96 species), and their parasitoids (126 species). Given its
312 replicated webs spanning large spatiotemporal scales, this dataset is well-suited for analyzing a variety of
313 ecological hypotheses and processes. Out of a total of 374 local webs, we retained those containing at least 5
314 species, resulting in a set of 233 georeferenced local webs (networks sampled within areas of 0.1 to 0.3 km²
315 during June and/or July spanning 29 years). We built a metaweb of binary interactions by aggregating all local
316 interactions, which gave us a regional web composed of 274 species and 1080 interactions. In the first two
317 panels of fig. 1, we show how the dissimilarity of interactions between common species (β_{OS}) and the
318 dissimilarity in species composition (β_S) between the metaweb and aggregated local webs (Poisot *et al.* (2012))
319 vary with the number of sampled local webs. This shows that networks of local interactions are highly
320 dissimilar from the metaweb, both in terms of species and interactions, especially when only a limited number
321 of sites has been sampled. Both dissimilarity indices were calculated based on the number of items shared by
322 the two webs (c_{LM}) and the number of items unique to the metaweb (u_M) and to the aggregated local web (u_L).
323 The β_S dissimilarity index uses species (nodes) as items being compared, while the β_{OS} index assesses

324 dissimilarity based on interactions between shared species (Poisot *et al.* (2012)). Both indices were calculated
325 following the β_W index of Whittaker (1960):

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

326 [Figure 1 about here.]

327 We converted these binary interaction networks into probabilistic ones using models based on simple
328 assumptions. Our models are not designed to estimate the exact values of probabilistic interactions. Instead,
329 their purpose is to create plausible networks that serve as illustrative examples to highlight distinctions between
330 local and metawebs of probabilistic interactions. We created two metawebs of probabilistic interactions by
331 employing constant false positive and false negative rates for all regional interactions. In the first metaweb, we
332 set both false positive and false negative rates to zero to prevent artificially inflating the total number of links,
333 enabling a more accurate comparison with binary interaction webs. This gave us a probability of regional
334 interaction of 1 when at least one interaction has been observed and of 0 in the absence of any observed
335 interaction between a given pair of species. In the second metaweb, we introduced a 5% false positive rate to
336 account for spurious interactions and a 10% false negative rate to address the elevated occurrence of missing
337 interactions in ecological networks (Catchen *et al.* (2023)). We believe these rates represent reasonable
338 estimates of missing and spurious potential interactions, but confirming their accuracy is challenging due to the
339 unavailability of data on the actual feasibility of interaction. Observed interactions were thus given a probability
340 of regional interaction of 95%, whereas unobserved ones were assigned a probability of 10%.

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

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

346 We built local webs of probabilistic interactions from the binary ones by using the metawebs of probabilistic
347 interactions and a constant value of $P(L|M)$ across interactions. Low values of $P(L|M)$ indicate that feasible

348 interactions rarely occur locally, intermediate values around 50% suggest considerable spatiotemporal
349 variability, while high values indicate that regional interactions are nearly always realized locally. Following
350 eq. 5, the local probability of interaction between a given pair of taxa consistently remained equal to or below
351 their probability of regional interaction.

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

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

358 where $P(L_1|M)$ and $P(L_2|M)$ are the probabilities of local interaction among two potentially interacting species
359 in the subnetworks L_1 and L_2 , respectively. This equation provides the probability that the interaction is realized
360 in either (1) exclusively the local web L_1 , (2) exclusively the local web L_2 or (3) both local webs, given that the
361 two taxa have the biological capacity to interact.

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

367 Spatial and temporal scaling of interactions

368 The investigation of network-area relationships and interaction accumulation curves is an important area of
369 research in network ecology. First, network-area relationships elucidate the scaling of network properties (such
370 as modularity and trophic chain lengths) with spatial scale (e.g., Wood *et al.* (2015), Galiana *et al.* (2018)). The
371 variations in network structure across spatial scales may stem from the scaling of species richness (species-area
372 relationships, SARs) and the number of interactions (Brose *et al.* (2004)) with the sampled area. Additionally,
373 ecological processes unfolding at distinct spatial scales, such as the spatial variability in local community

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

385 Local and metawebs exhibit distinct relationships with spatial and temporal scales. On one hand, metawebs of
386 probabilistic interactions, being independent of any local context, feature regional interactions that do not scale
387 with space and time. This is because regional interactions depend solely on the biological capacity of two taxa
388 to interact, regardless of their co-occurrence and specific environmental conditions. However, probabilities of
389 potential interactions may change (tending to become more definitive) upon updating previous estimates with
390 increased sampling effort, even though they do not vary in a specific direction with the spatial and temporal
391 extent (boundaries) of the network. The probability of two taxa potentially interacting should theoretically be
392 the same in all metawebs in which they are present, provided that the data and methods used for estimation are
393 consistent. For example, if a smaller metaweb M_1 is derived from a larger metaweb M_0 by selecting the subset
394 of taxa present in the region described by M_1 and retaining all their interactions, their probabilities of
395 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
396 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
397 larger web, even though pairwise probabilities remain identical.

398 On the other hand, local interactions scale both spatially and temporally, given that they have more
399 opportunities to be realized and observed in larger areas and longer durations. This is attributed to factors such
400 as a higher number of individuals, greater trait variations, and increased opportunities for encounters, as
401 highlighted by McLeod *et al.* (2020). For example, if a local web of probabilistic interactions L_1 with an area
402 A_1 is derived from a larger web L_0 with an area A_0 , and A_1 is entirely nested within A_0 , interaction probabilities
403 should be 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 probabilities could be higher in the smaller area, contingent upon their environmental and biological conditions. Likewise, interaction probabilities are expected to be lower in webs with shorter durations when time intervals are nested. In fig. 2, we show how the expected number of local host-parasite interactions scales with area, represented as an expanding latitudinal window, in comparison with regional interactions. Even though we employed local probabilities of interactions equal to regional interactions for the purpose of comparison (i.e., using $P(L|M) = 1$ here), we notice that the total number of regional interactions scales more rapidly than local interactions. This is because numerous regional interactions involve species that never co-occur, and as a result, these interactions are not captured in local webs.

[Figure 2 about here.]

Box 1: A spatiotemporally explicit model of interactions

Predicting local webs across time and space is a pivotal goal of network ecology, especially given the scarcity of interaction data (Strydom *et al.* (2021)). Ecologists may resort to predictive models (e.g., generative Bayesian and machine learning models) to reconstruct networks at fine spatial and temporal scales with limited interaction data. For instance, real-time biomonitoring data coupled with appropriate numerical models (Bohan *et al.* (2017)) can be employed to reconstruct local ecological webs, opening avenues for in-depth studies on local ecosystem functioning and dynamics. The probabilistic representation of interactions acknowledges the inherent uncertainty in these models, typically expressed through probability distributions. We introduce and develop a simple generative mechanistic model for probabilistic local interactions that takes into consideration their spatiotemporal variability (i.e. a spatiotemporally explicit model of local interactions). It is essential to note that our model is not designed for regional interactions, which are scale-independent. Rather, it could prove valuable for predicting local interactions across time and space by generating new interaction data following parameter inference.

As indicated by eq. 1, the probability that two taxa i and j interact at a specific location (x, y, z) may be determined by the product of their probability of interaction given co-occurrence and their probability of co-occurrence. First, their co-occurrence probability can be calculated using their individual occurrence probabilities $P(X_i)$ and $P(X_j)$, along with the strength of association γ between their occurrences and co-occurrence (Cazelles *et al.* (2016)):

$$P(C_{i,j}) = P(X_i, X_j) = P(X_i)P(X_j)\gamma. \quad (7)$$

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

$$C_{i,j} \sim \text{Bernoulli}(P(X_i, X_j)). \quad (8)$$

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

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

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

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

$$L_{i \rightarrow j} \sim \text{Bernoulli}(P(L_{i \rightarrow j})) \quad (10)$$

$$P(L_{i \rightarrow j}) = P(X_i)P(X_j)\gamma(1 - e^{-\lambda t_0}) \quad (11)$$

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

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

449 In fig. 3, we show the variation in the probability of interaction under different parameter values. In the right
 450 panel, we notice that, irrespective of the interaction rate λ , the probability of interaction converges toward an
 451 asymptote determined by the co-occurrence $P(C_{i,j})$ (eq. 7). This model can be customized in different ways,
 452 such as by linking λ with specific environmental variables or explicitly incorporating observation errors (i.e.,
 453 the probabilities of false negatives and false positives).

454 [Figure 3 about here.]

455 Taxonomic scaling of interactions

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

464 There are no inherent differences in the taxonomic scaling between local and metawebs, as only the nodes are
 465 defined taxonomically. In other words, the probability values of edges in both local (eq. 2) and metawebs (eq. 3)
 466 are not conditional on any taxonomic scale. The phylogenetic scale is tied to the definition of the event itself
 467 (i.e., the interaction between two taxa, defined at the desired taxonomical scale), not to the variables on which
 468 interaction probabilities are conditional. In both types of webs, transitioning to a broader level of organization

469 (e.g., from a species-level web S to a genus-level web G) can be accomplished directly by using probabilities
 470 from finer scales. For example, in a network with n_1 species from genus g_1 and n_2 species from genus g_2 , one
 471 can compute the probability that at least one species from genus g_1 interacts with at least one species from
 472 genus g_2 as follows:

$$P(G_{g_1 \rightarrow g_2}) = 1 - \prod_{i=1}^{n_1} \prod_{j=1}^{n_2} (1 - P(S_{g_{1i} \rightarrow g_{2j}})), \quad (14)$$

473 where g_{1i} and g_{2j} are the species of the corresponding genus. If it is known that at least two of these species
 474 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
 475 equal to 1. Canard *et al.* (2012) built a species-based network following a similar approach, by using simulated
 476 interactions between individuals derived from a neutral model (i.e., a model that assumed ecological
 477 equivalence among individuals). In contrast, a more sophisticated approach is necessary when transitioning
 478 from a broader to a finer level of organization. This is because knowledge of an interaction between two genera
 479 does not guarantee that all possible pairwise combinations of their species will also interact. One possible
 480 method is to build a finer-scale network by generating probabilities of interactions through random sampling
 481 from a beta distribution, parameterized by the broader-scale network.

482 Ideally, our biological interpretation of probabilistic interactions should remain consistent across a network
 483 even if it incorporates heterogeneous levels of organization, such as a network whose nodes represent both
 484 species and trophic species (groups of species sharing similar predators and prey). This is common in
 485 ecological webs where taxonomic resolution is typically low (Hemprich-Bennett *et al.* (2021), Vázquez§ *et al.*
 486 (2022)). Interaction probabilities at broader taxonomic scales should be based on probabilities of interactions
 487 between individuals, for both local and metawebs. For instance, in local individual-based food webs, the
 488 probability that two individuals interact reflects our degree of belief that one individual will consume the other.
 489 Similarly, in local species-based food webs, the probability that two species interact represents our degree of
 490 belief that at least one individual from the predator species consumes at least another individual from the prey
 491 species. Furthermore, in local clade-based food webs, the probability that two clades interact represents our
 492 degree of belief that at least two species from these clades interact with each other or, equivalently, that at least
 493 two individuals from these clades interact with each other. Fundamentally, the taxonomic scaling of interactions
 494 involves aggregating interactions between individuals into larger groups, which may exhibit varying degrees of
 495 homogeneity based on the organisms and the study system. In that regard, taxonomic scaling is analogous to the

496 spatial and temporal scaling of interactions, as they all represent different ways to aggregate individuals into
497 broader groups (either spatially, temporally, or taxonomically).

498 **Sampling for binary interaction webs**

499 The prediction of binary interactions through Bernoulli trials is an important application of probabilistic
500 interaction webs. This approach proves beneficial for analyzing the structural characteristics of probabilistic
501 interaction webs, particularly in the absence of specific analytical measures. By performing independent
502 Bernoulli trials for each interaction, a network of binary interactions may be generated. A probability
503 distribution of network properties can be obtained by measuring network structure across multiple randomly
504 generated networks (Poisot *et al.* (2016)). This method enables the representation of the variability of network
505 structure, albeit with possible biases when connectance is low (Poisot & Gravel (2014), Chagnon (2015)).
506 Employing this strategy to generate binary interaction networks under a null model facilitates null hypothesis
507 significance testing, wherein the observed measure is compared against the simulated distribution (e.g.,
508 Bascompte *et al.* (2003)). Additionally, the random generation of binary interaction networks, from a
509 probabilistic interaction web that accounts for the spatiotemporal variability of interactions, may effectively
510 capture network structure across space and time. This facilitates the investigation of ecological hypotheses
511 about interactions at broad spatial and temporal scales.

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

522 In fig. 4, we compare the average connectance of binary interaction networks resulting from these two sampling
523 techniques, where regional and local interactions are drawn from our host-parasite networks of probabilistic

interactions, generating a number of binary interaction web realizations for each site in the dataset. These two sampling approaches yield different outcomes, particularly for lower values of $P(L|M)$, which denote instances where regional interactions do not consistently manifest locally. Small discrepancies are also apparent between these techniques when we equate the probability of local interaction to the probability of regional interaction (i.e., when using $P(L|M) = 1.0$ in eq. 5), especially when the number of samples of binary interaction networks for each location is low. As anticipated, we observe that sampling binary interactions from the metaweb tends to overestimate connectance on average compared to sampling them from local webs. Furthermore, we observe an increase in the variability of connectance when employing a single sample, representing what we consider as a more tangible process leading to the realization of local and regional interactions in nature.

[Figure 4 about here.]

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

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

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

Prediction of local webs from metawebs

Metawebs serve as a valuable source of ecological information for predicting local webs across time and space. Local webs of binary interactions can be reconstructed by selecting a subset of taxa and interactions from the metaweb (Dunne (2006)). This implies that metawebs consistently contain more interactions than their corresponding local webs. Determining the list of taxa to select can be achieved empirically or through methods

549 like range maps or species distribution models. As the species composition of a community is arguably less
550 difficult to sample or predict than its interactions, the primary challenge lies in deciding which interactions to
551 select from the metaweb, a task that may necessitate advanced statistical models and ecological expertise.
552 Inferring the structure of local webs from the metaweb before predicting specific local pairwise interactions
553 could hold promise (Strydom *et al.* (2021)), considering that the structure of local webs is constrained by the
554 metaweb (Saravia *et al.* (2022)).

555 Building local webs of probabilistic interactions from a metaweb of probabilistic interactions involves a
556 reduction in the value of pairwise interaction probabilities. This decrease is attributed to the prerequisite that
557 two taxa must initially possess the capacity to interact before engaging in local interactions (eq. 5). Therefore,
558 inferring local webs from their metaweb while maintaining identical interaction probability values would
559 introduce systematic biases into the predictions. In such cases, these networks would essentially represent
560 smaller-scale metawebs of potential interactions, possibly leading to misinterpretations by being perceived as
561 local interactions. As proposed by McLeod *et al.* (2021), although metawebs do not capture the spatiotemporal
562 variability of interactions, they establish an upper limit for local interactions. In other words, the probability of
563 two taxa interacting at a specific location and time is consistently lower or equal to the probability of their
564 regional interaction:

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

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

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

569 Estimating more precisely the probability $P(L|M)$ that two taxa interact locally if they can potentially interact
570 allows for improved predictions of local webs from a metaweb of probabilistic interactions. This task is
571 challenging due to the variability of this probability across space and time, as well as its variability across
572 pairwise interactions within a network. Using simple models of $P(L|M)$, as demonstrated in our case studies,

573 represents an initial step toward the overarching objective of reconstructing local webs from metawebs.

574 Conclusion

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

587 It is essential to enhance our comprehension of both regional and local interactions, especially considering the
588 current scarcity of interaction data. However, while sampling biological communities does decrease the
589 uncertainty of interactions by accumulating evidence for their feasibility and local realization, there is a limit to
590 how much we can diminish their inherent randomness. In metawebs, probabilities reflect our limited knowledge
591 of interactions (i.e., our degree of belief that interactions are feasible), which is expected to improve with a
592 larger volume of data. We should anticipate that regional interactions will become more definitive (with
593 probabilities approaching 0 or 1) as we investigate various conditions, including different combinations of
594 species traits. Conversely, in the case of local webs, which can be seen as random instances of metawebs,
595 randomness cannot be as reduced. In local interactions, probabilities can be divided into a component
596 representing uncertainty and another representing spatiotemporal variability. Owing to environmental
597 heterogeneity, there will invariably be instances in which an interaction occurs and others in which it does not,
598 across different times and locations, irrespective of the extent to which we can improve our knowledge of its
599 biological feasibility and the local conditions that facilitate its occurrence. When local webs depict probabilities
600 of observing interactions rather than just their actual occurrence, we must also consider the observation

variability as an additional source of randomness. Every ecological process is stochastic but there is also a possibility that a phenomenon goes undetected. Quantifying and partitioning this stochasticity will enable us to make more accurate predictions about ecological interactions at various spatial and temporal scales. This will prove to be of vital importance as our time to understand nature runs out, especially at the places where the impacts of climate change and habitat loss hit harder.

Acknowledgment

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

References

- Abrego, N., Roslin, T., Huotari, T., Ji, Y., Schmidt, N.M., Wang, J., *et al.* (2021). Accounting for species interactions is necessary for predicting how arctic arthropod communities respond to climate change. *Ecography*, 44, 885–896.
- Andrade-Pacheco, R., Rerolle, F., Lemoine, J., Hernandez, L., Meïté, A., Juziwelo, L., *et al.* (2020). Finding hotspots: Development of an adaptive spatial sampling approach. *Scientific Reports*, 10, 10939.
- Angilletta, M.J., Jr., Steury, T.D. & Sears, M.W. (2004). Temperature, Growth Rate, and Body Size in Ectotherms: Fitting Pieces of a Life-History Puzzle¹. *Integrative and Comparative Biology*, 44, 498–509.
- Araujo, M.B., Rozenfeld, A., Rahbek, C. & Marquet, P.A. (2011). Using species co-occurrence networks to assess the impacts of climate change. *Ecography*, 34, 897–908.
- Banašek-Richter, C., Cattin, M.-F. & Bersier, L.-F. (2004). Sampling effects and the robustness of quantitative and qualitative food-web descriptors. *Journal of Theoretical Biology*, 226, 23–32.

- 626 Bartomeus, I., Gravel, D., Tylianakis, J.M., Aizen, M.A., Dickie, I.A. & Bernard-Verdier, M. (2016). [A common framework for identifying linkage rules across different types of interactions](#). *Functional Ecology*, 30, 1894–1903.
- 627
- 628
- 629 Bascompte, J., Jordano, P., Melián, C.J. & Olesen, J.M. (2003). [The nested assembly of plant–animal mutualistic networks](#). *Proceedings of the National Academy of Sciences*, 100, 9383–9387.
- 630
- 631 Berlow, E.L., Neutel, A.-M., Cohen, J.E., De Ruiter, P.C., Ebenman, B., Emmerson, M., *et al.* (2004). [Interaction strengths in food webs: Issues and opportunities](#). *Journal of Animal Ecology*, 73, 585–598.
- 632
- 633 Bersier, L.-F., Dixon, P. & Sugihara, G. (1999). [Scale-Invariant or Scale-Dependent Behavior of the Link Density Property in Food Webs: A Matter of Sampling Effort?](#) *The American Naturalist*, 153, 676–682.
- 634
- 635 Blanchet, F.G., Cazelles, K. & Gravel, D. (2020). [Co-occurrence is not evidence of ecological interactions](#). *Ecology Letters*, 23, 1050–1063.
- 636
- 637 Bohan, D.A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A.J. & Woodward, G. (2017). [Next-Generation Global Biomonitoring: Large-scale, Automated Reconstruction of Ecological Networks](#). *Trends in Ecology & Evolution*, 32, 477–487.
- 638
- 639
- 640 Bolnick, D.I., Amarasekare, P., Araújo, M.S., Bürger, R., Levine, J.M., Novak, M., *et al.* (2011). [Why intraspecific trait variation matters in community ecology](#). *Trends in Ecology & Evolution*, 26, 183–192.
- 641
- 642 Borrett, S.R. & Scharler, U.M. (2019). [Walk partitions of flow in Ecological Network Analysis: Review and synthesis of methods and indicators](#). *Ecological Indicators*, 106, 105451.
- 643
- 644 Boulangeat, I., Gravel, D. & Thuiller, W. (2012). [Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances](#). *Ecology Letters*, 15, 584–593.
- 645
- 646 Brimacombe, C., Bodner, K., Michalska-Smith, M., Poisot, T. & Fortin, M.-J. (2023). [Shortcomings of reusing species interaction networks created by different sets of researchers](#). *PLOS Biology*, 21, e3002068.
- 647
- 648 Broom, M. & Ruxton, G.D. (2005). [You can run—or you can hide: Optimal strategies for cryptic prey against pursuit predators](#). *Behavioral Ecology*, 16, 534–540.
- 649
- 650 Brose, U., Ostling, A., Harrison, K. & Martinez, N.D. (2004). [Unified spatial scaling of species and their trophic interactions](#). *Nature*, 428, 167–171.
- 651
- 652 Canard, E.F., Mouquet, N., Mouillot, D., Stanko, M., Miklisova, D. & Gravel, D. (2014). [Empirical Evaluation of Neutral Interactions in Host-Parasite Networks](#). *The American Naturalist*, 183, 468–479.
- 653

- 654 Canard, E., Mouquet, N., Marescot, L., Gaston, K.J., Gravel, D. & Mouillot, D. (2012). [Emergence of](#)
655 [Structural Patterns in Neutral Trophic Networks.](#) *PLOS ONE*, 7, e38295.
- 656 Catchen, M.D., Poisot, T., Pollock, L.J. & Gonzalez, A. (2023). The missing link: Discerning true from false
657 negatives when sampling species interaction networks.
- 658 Cazelles, K., Araujo, M.B., Mouquet, N. & Gravel, D. (2016). [A theory for species co-occurrence in interaction](#)
659 [networks.](#) *Theoretical Ecology*, 9, 39–48.
- 660 Chagnon, P.-L. (2015). [Characterizing topology of ecological networks along gradients: The limits of metrics'](#)
661 [standardization.](#) *Ecological Complexity*, 22, 36–39.
- 662 Choh, Y., Ignacio, M., Sabelis, M.W. & Janssen, A. (2012). [Predator-prey role reversals, juvenile experience](#)
663 [and adult antipredator behaviour.](#) *Scientific Reports*, 2, 728.
- 664 Cirtwill, A.R., Eklöf, A., Roslin, T., Wootton, K. & Gravel, D. (2019). [A quantitative framework for](#)
665 [investigating the reliability of empirical network construction.](#) *Methods in Ecology and Evolution*, 10,
666 902–911.
- 667 Dallas, T., Park, A.W. & Drake, J.M. (2017). [Predicting cryptic links in host-parasite networks.](#) *PLOS*
668 [Computational Biology, 13, e1005557.](#)
- 669 Dunne, J.A. (2006). The Network Structure of Food Webs. In: *Ecological networks: Linking structure and*
670 *dynamics* (eds. Dunne, J.A. & Pascual, M.). Oxford University Press, pp. 27–86.
- 671 Elmasri, M., Farrell, M.J., Davies, T.J. & Stephens, D.A. (2020). [A hierarchical Bayesian model for predicting](#)
672 [ecological interactions using scaled evolutionary relationships.](#) *The Annals of Applied Statistics*, 14,
673 221–240.
- 674 Elton, C.S. (2001). *Animal Ecology*. University of Chicago Press, Chicago, IL.
- 675 Emmerson, M.C. & Raffaelli, D. (2004). [Predator–prey body size, interaction strength and the stability of a real](#)
676 [food web.](#) *Journal of Animal Ecology*, 73, 399–409.
- 677 Fortuna, M.A. & Bascompte, J. (2006). [Habitat loss and the structure of plant–animal mutualistic networks.](#)
678 *Ecology Letters*, 9, 281–286.
- 679 Fu, X., Seo, E., Clarke, J. & Hutchinson, R.A. (2021). [Link Prediction Under Imperfect Detection:](#)
680 [Collaborative Filtering for Ecological Networks.](#) *IEEE Transactions on Knowledge and Data Engineering*,
681 33, 3117–3128.

- 682 Galiana, N., Lurgi, M., Claramunt-López, B., Fortin, M.-J., Leroux, S., Cazelles, K., *et al.* (2018). [The spatial](#)
683 [scaling of species interaction networks](#). *Nature Ecology & Evolution*, 2, 782–790.
- 684 Gilman, S.E., Urban, M.C., Tewksbury, J., Gilchrist, G.W. & Holt, R.D. (2010). [A framework for community](#)
685 [interactions under climate change](#). *Trends in Ecology & Evolution*, 25, 325–331.
- 686 Godsoe, W., Murray, R. & Iritani, R. (2022). [Species interactions and diversity: A unified framework using Hill](#)
687 [numbers](#). *Oikos*, n/a, e09282.
- 688 Golubski, A.J. & Abrams, P.A. (2011). [Modifying modifiers: What happens when interspecific interactions](#)
689 [interact?](#) *Journal of Animal Ecology*, 80, 1097–1108.
- 690 Gomez, J.M., Verdu, M. & Perfectti, F. (2010). [Ecological interactions are evolutionarily conserved across the](#)
691 [entire tree of life](#). *Nature*, 465, 918–U6.
- 692 Gonzalez, A. & Londoño, M.C. (2022). [Monitor biodiversity for action](#). *Science*, 378, 1147–1147.
- 693 Gonzalez-Varo, J.P. & Traveset, A. (2016). [The Labile Limits of Forbidden Interactions](#). *Trends in Ecology &*
694 *Evolution*, 31, 700–710.
- 695 Gravel, D., Albouy, C. & Thuiller, W. (2016). [The meaning of functional trait composition of food webs for](#)
696 [ecosystem functioning](#). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371,
697 20150268.
- 698 Gravel, D., Baiser, B., Dunne, J.A., Kopelke, J.-P., Martinez, N.D., Nyman, T., *et al.* (2019). [Bringing Elton](#)
699 [and Grinnell together: A quantitative framework to represent the biogeography of ecological interaction](#)
700 [networks](#). *Ecography*, 42, 401–415.
- 701 Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. (2013). [Inferring food web structure from](#)
702 [predator–prey body size relationships](#). *Methods in Ecology and Evolution*, 4, 1083–1090.
- 703 Guimarães, P.R. (2020). [The Structure of Ecological Networks Across Levels of Organization](#). *Annual Review*
704 *of Ecology, Evolution, and Systematics*, 51, 433–460.
- 705 Guimerà, R. & Sales-Pardo, M. (2009). [Missing and spurious interactions and the reconstruction of complex](#)
706 [networks](#). *Proceedings of the National Academy of Sciences*, 106, 22073–22078.
- 707 Hemprich-Bennett, D.R., Oliveira, H.F.M., Le Comber, S.C., Rossiter, S.J. & Clare, E.L. (2021). [Assessing the](#)
708 [impact of taxon resolution on network structure](#). *Ecology*, 102, e03256.

- 709 Herrera, C.M. (1989). [Pollinator abundance, morphology, and flower visitation rate: Analysis of the “quantity”](#)
710 [component in a plant-pollinator system.](#) *Oecologia*, 80, 241–248.
- 711 Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M. & Ladle, R.J. (2015). [Seven Shortfalls](#)
712 [that Beset Large-Scale Knowledge of Biodiversity.](#) *Annual Review of Ecology, Evolution, and Systematics*,
713 46, 523–549.
- 714 Jordano, P. (1987). [Patterns of Mutualistic Interactions in Pollination and Seed Dispersal: Connectance,](#)
715 [Dependence Asymmetries, and Coevolution.](#) *The American Naturalist*, 129, 657–677.
- 716 Jordano, P. (2016). [Sampling networks of ecological interactions.](#) *Functional Ecology*, 30, 1883–1893.
- 717 Jordano, P., Bascompte, J. & Olesen, J.M. (2003). [Invariant properties in coevolutionary networks of](#)
718 [plant–animal interactions.](#) *Ecology Letters*, 6, 69–81.
- 719 Kéfi, S., Berlow, E.L., Wieters, E.A., Joppa, L.N., Wood, S.A., Brose, U., *et al.* (2015). [Network structure](#)
720 [beyond food webs: Mapping non-trophic and trophic interactions on Chilean rocky shores.](#) *Ecology*, 96,
721 291–303.
- 722 Kéfi, S., Berlow, E.L., Wieters, E.A., Navarrete, S.A., Petchey, O.L., Wood, S.A., *et al.* (2012). [More than a](#)
723 [meal... integrating non-feeding interactions into food webs.](#) *Ecology Letters*, 15, 291–300.
- 724 Kéfi, S., Miele, V., Wieters, E.A., Navarrete, S.A. & Berlow, E.L. (2016). [How Structured Is the Entangled](#)
725 [Bank? The Surprisingly Simple Organization of Multiplex Ecological Networks Leads to Increased](#)
726 [Persistence and Resilience.](#) *PLOS Biology*, 14, e1002527.
- 727 Klecka, J. & Boukal, D.S. (2014). [The effect of habitat structure on prey mortality depends on predator and](#)
728 [prey microhabitat use.](#) *Oecologia*, 176, 183–191.
- 729 Kopelke, J.-P., Nyman, T., Cazelles, K., Gravel, D., Vissault, S. & Roslin, T. (2017). [Food-web structure of](#)
730 [willow-galling sawflies and their natural enemies across Europe.](#) *Ecology*, 98, 1730–1730.
- 731 Kulkarni, M.A., Kweka, E., Nyale, E., Lyatuu, E., Mosha, F.W., Chandramohan, D., *et al.* (2006).
732 [Entomological Evaluation of Malaria Vectors at Different Altitudes in Hai District, Northeastern Tanzania.](#)
733 [Journal of Medical Entomology, 43, 580–588.](#)
- 734 McCann, K. (2007). Protecting biostructure. *Nature*, 446, 29–29.
- 735 McCann, K.S. (2011). Food webs (MPB-50). In: *Food Webs (MPB-50)*. Princeton University Press.

- 736 McLeod, A., Leroux, S.J., Gravel, D., Chu, C., Cirtwill, A.R., Fortin, M.-J., *et al.* (2021). [Sampling and](#)
737 [asymptotic network properties of spatial multi-trophic networks](#). *Oikos*, n/a.
- 738 McLeod, A.M., Leroux, S.J. & Chu, C. (2020). [Effects of species traits, motif profiles, and environment on](#)
739 [spatial variation in multi-trophic antagonistic networks](#). *Ecosphere*, 11, e03018.
- 740 Melián, C.J., Vilas, C., Baldó, F., González-Ortegón, E., Drake, P. & Williams, R.J. (2011). [Eco-evolutionary](#)
741 [Dynamics of Individual-Based Food Webs](#). In: *Advances in Ecological Research*, The Role of Body Size in
742 Multispecies Systems (ed. Belgrano, A.). Academic Press, pp. 225–268.
- 743 Ohlmann, M., Miele, V., Dray, S., Chalmandrier, L., O'Connor, L. & Thuiller, W. (2019). [Diversity indices for](#)
744 [ecological networks: A unifying framework using Hill numbers](#). *Ecology Letters*, 22, 737–747.
- 745 Olesen, J.M., Bascompte, J., Dupont, Y.L., Elberling, H., Rasmussen, C. & Jordano, P. (2010). [Missing and](#)
746 [forbidden links in mutualistic networks](#). *Proceedings of the Royal Society B: Biological Sciences*, 278,
747 725–732.
- 748 Pascual, M., Dunne, J.A. & Dunne, J.A. (2006). *Ecological Networks: Linking Structure to Dynamics in Food*
749 [Webs](#). Oxford University Press, USA.
- 750 Petchey, O.L., Beckerman, A.P., Riede, J.O. & Warren, P.H. (2008). [Size, foraging, and food web structure](#).
751 *Proceedings of the National Academy of Sciences*, 105, 4191–4196.
- 752 Pilosof, S., Porter, M.A., Pascual, M. & Kéfi, S. (2017). [The multilayer nature of ecological networks](#). *Nature*
753 [Ecology & Evolution](#), 1, 1–9.
- 754 Plagányi, É.E. & Butterworth, D.S. (2004). [A critical look at the potential of Ecopath with ecosim to assist in](#)
755 [practical fisheries management](#). *African Journal of Marine Science*, 26, 261–287.
- 756 Poisot, T., Canard, E., Mouillot, D., Mouquet, N. & Gravel, D. (2012). [The dissimilarity of species interaction](#)
757 [networks](#). *Ecology Letters*, 15, 1353–1361.
- 758 Poisot, T., Cirtwill, A.R., Cazelles, K., Gravel, D., Fortin, M.-J. & Stouffer, D.B. (2016). [The structure of](#)
759 [probabilistic networks](#). *Methods in Ecology and Evolution*, 7, 303–312.
- 760 Poisot, T. & Gravel, D. (2014). [When is an ecological network complex? Connectance drives degree](#)
761 [distribution and emerging network properties](#). *PeerJ*, 2, e251.
- 762 Poisot, T., Stouffer, D.B. & Gravel, D. (2015). [Beyond species: Why ecological interaction networks vary](#)
763 [through space and time](#). *Oikos*, 124, 243–251.

- 764 Polis, G.A. (1991). [Complex Trophic Interactions in Deserts: An Empirical Critique of Food-Web Theory](#). *The
765 American Naturalist*, 138, 123–155.
- 766 Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O’Hara, R.B., Parris, K.M., *et al.* (2014). [Understanding
767 co-occurrence by modelling species simultaneously with a Joint Species Distribution Model \(JSDM\)](#).
768 *Methods in Ecology and Evolution*, 5, 397–406.
- 769 Pomeranz, J.P.F., Thompson, R.M., Poisot, T. & Harding, J.S. (2019). [Inferring predator–prey interactions in
770 food webs](#). *Methods in Ecology and Evolution*, 10, 356–367.
- 771 Pringle, R.M. & Hutchinson, M.C. (2020). [Resolving Food-Web Structure](#). *Annual Review of Ecology,
772 Evolution, and Systematics*, 51, 55–80.
- 773 Proulx, S.R., Promislow, D.E.L. & Phillips, P.C. (2005). [Network thinking in ecology and evolution](#). *Trends in
774 Ecology & Evolution*, SPECIAL ISSUE: BUMPER BOOK REVIEW, 20, 345–353.
- 775 Pulliam, H.R. (1974). [On the Theory of Optimal Diets](#). *The American Naturalist*, 108, 59–74.
- 776 Rooney, N. & McCann, K.S. (2012). [Integrating food web diversity, structure and stability](#). *Trends in Ecology
777 & Evolution*, 27, 40–46.
- 778 Salim, J.A., Saraiva, A.M., Zermoglio, P.F., Agostini, K., Wolowski, M., Drucker, D.P., *et al.* (2022). [Data
779 standardization of plant–pollinator interactions](#). *GigaScience*, 11, giac043.
- 780 Sanders, D. & van Veen, F.J.F. (2012). [Indirect commensalism promotes persistence of secondary consumer
781 species](#). *Biology Letters*, 8, 960–963.
- 782 Saravia, L.A., Marina, T.I., Kristensen, N.P., De Troch, M. & Momo, F.R. (2022). [Ecological network
783 assembly: How the regional metaweb influences local food webs](#). *Journal of Animal Ecology*, 91, 630–642.
- 784 Schmeller, D.S., Julliard, R., Bellingham, P.J., Böhm, M., Brummitt, N., Chiarucci, A., *et al.* (2015). [Towards a
785 global terrestrial species monitoring program](#). *Journal for Nature Conservation*, 25, 51–57.
- 786 Shaw, J.O., Dunhill, A.M., Beckerman, A.P., Dunne, J.A. & Hull, P.M. (2024). [A framework for reconstructing
787 ancient food webs using functional trait data](#).
- 788 Singer, M.C. & McBride, C.S. (2012). [Geographic mosaics of species’ association: A definition and an
789 example driven by plant–insect phenological synchrony](#). *Ecology*, 93, 2658–2673.

- 790 Stouffer, D.B., Rezende, E.L. & Amaral, L.A.N. (2011). [The role of body mass in diet contiguity and food-web](#)
791 [structure](#). *Journal of Animal Ecology*, 80, 632–639.
- 792 Strydom, T., Bouskila, S., Banville, F., Barros, C., Caron, D., Farrell, M.J., *et al.* (2022). [Food web](#)
793 [reconstruction through phylogenetic transfer of low-rank network representation](#). *Methods in Ecology and*
794 *Evolution*, 13.
- 795 Strydom, T., Catchen, M.D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., *et al.* (2021). [A](#)
796 [roadmap towards predicting species interaction networks \(across space and time\)](#). *Philosophical*
797 *Transactions of the Royal Society B-Biological Sciences*, 376, 20210063.
- 798 Tylianakis, J.M., Tscharntke, T. & Lewis, O.T. (2007). [Habitat modification alters the structure of tropical](#)
799 [host-parasitoid food webs](#). *Nature*, 445, 202–205.
- 800 Vázquez§, D.P., Peralta§, G., Cagnolo, L., Santos, M. & Igual, §.E. autores contribuyeron por. (2022).
801 [Ecological interaction networks. What we know, what we don't, and why it matters](#). *Ecología Austral*, 32,
802 670–697.
- 803 Weinstein, B.G. & Graham, C.H. (2017a). [On comparing traits and abundance for predicting species](#)
804 [interactions with imperfect detection](#). *Food Webs*, 11, 17–25.
- 805 Weinstein, B.G. & Graham, C.H. (2017b). [Persistent bill and corolla matching despite shifting temporal](#)
806 [resources in tropical hummingbird-plant interactions](#). *Ecology Letters*, 20, 326–335.
- 807 Whittaker, R.H. (1960). [Vegetation of the Siskiyou Mountains, Oregon and California](#). *Ecological*
808 *Monographs*, 30, 279–338.
- 809 Wood, S.A., Russell, R., Hanson, D., Williams, R.J. & Dunne, J.A. (2015). [Effects of spatial scale of sampling](#)
810 [on food web structure](#). *Ecology and Evolution*, 5, 3769–3782.
- 811 Woodward, G., Benstead, J.P., Beveridge, O.S., Blanchard, J., Brey, T., Brown, L.E., *et al.* (2010). [Chapter 2 -](#)
812 [Ecological Networks in a Changing Climate](#). In: *Advances in Ecological Research*, Ecological Networks
813 (ed. Woodward, G.). Academic Press, pp. 71–138.
- 814 Woodward, G., Brown, L.E., Edwards, F.K., Hudson, L.N., Milner, A.M., Reuman, D.C., *et al.* (2012). [Climate](#)
815 [change impacts in multispecies systems: Drought alters food web size structure in a field experiment](#).
816 *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 2990–2997.

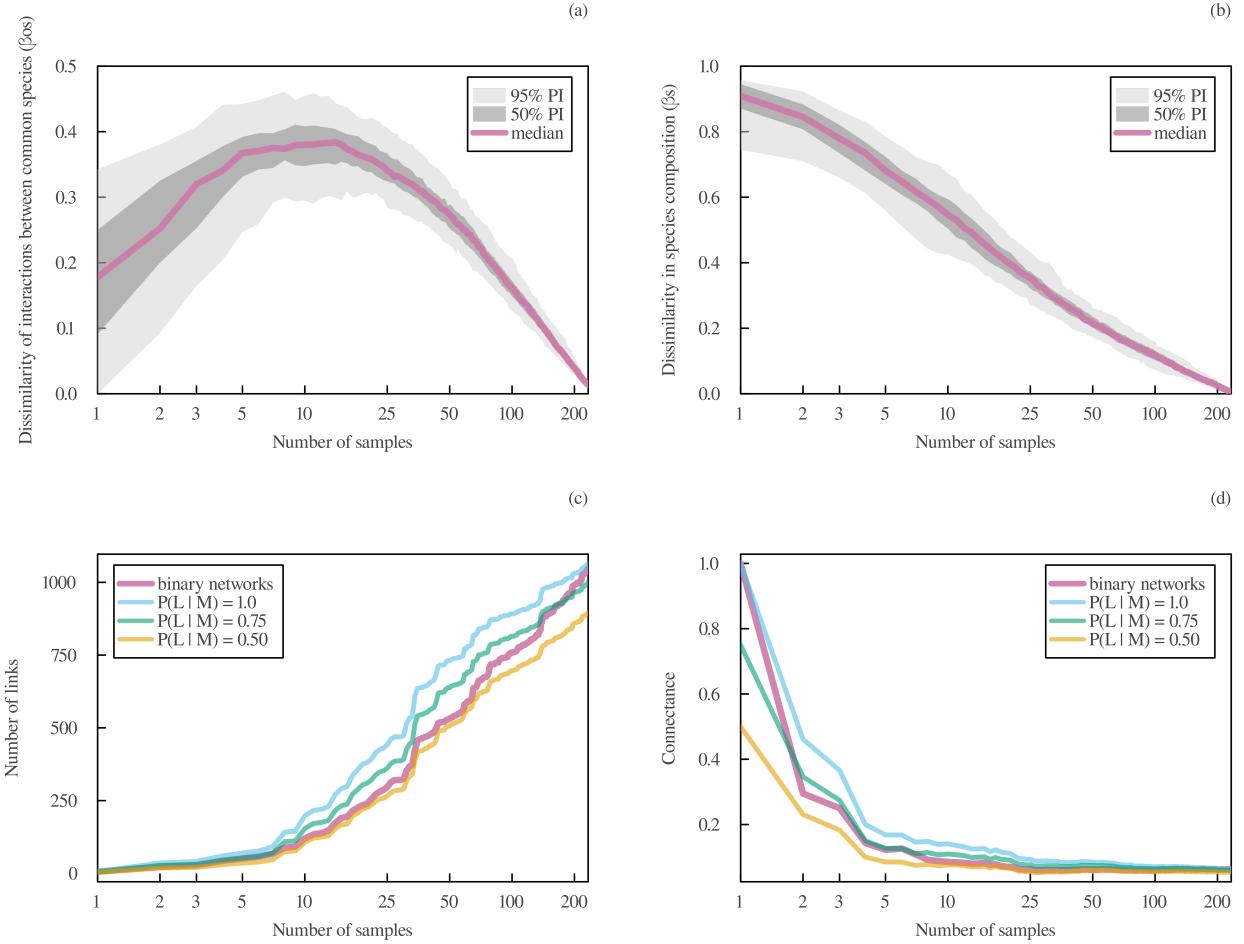


Figure 1: Network accumulation curves. (a) Dissimilarity of interactions between common species and (b) dissimilarity in species composition between aggregated local webs and the metaweb of binary host-parasite interactions. Aggregated local webs were obtained by sequentially and randomly selecting a number of local webs of binary interactions and aggregating both their species and interactions. In both panels, the colored line represents the median dissimilarity across simulations and the grey areas cover the 50% and 95% percentile intervals. (c) Scaling of the number of links and (d) scaling of connectance with the number of sampled binary and probabilistic local interaction webs. For a better comparison with binary interactions, local webs of probabilistic interactions were derived from the metaweb of probabilistic interactions 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 local webs of probabilistic interactions were obtained by sequentially and randomly selecting a number of local webs and aggregating both their species and interactions (with the value of $P(L|M)$ adjusting according to eq. 6).

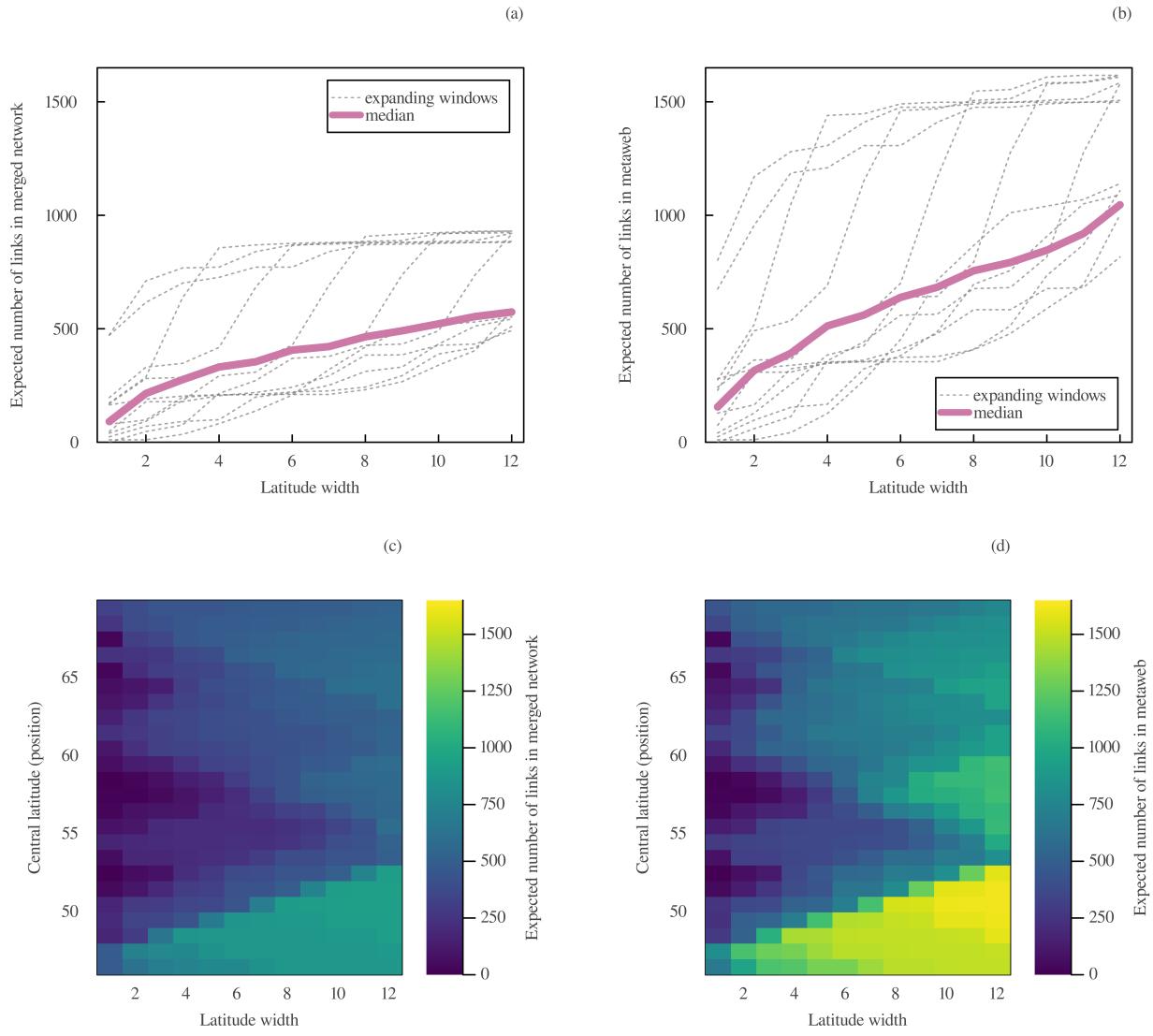


Figure 2: Spatial scaling of interactions. Expected number of host-parasite interactions in a network aggregating all (a) local and (b) regional probabilistic interactions within a latitudinal window of a given length. Every dashed curve corresponds to a different expanding window characterized by a different central latitude, with the colored solid line representing the median number of interactions. Heatmaps of the expected number of (c) local and (d) regional interactions found in windows of specified length and central latitudes. Probabilities of regional interactions were obtained with a false positive rate of 5% and a false negative rate of 10%. Local probabilistic interactions were derived from the probabilistic regional interactions 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.

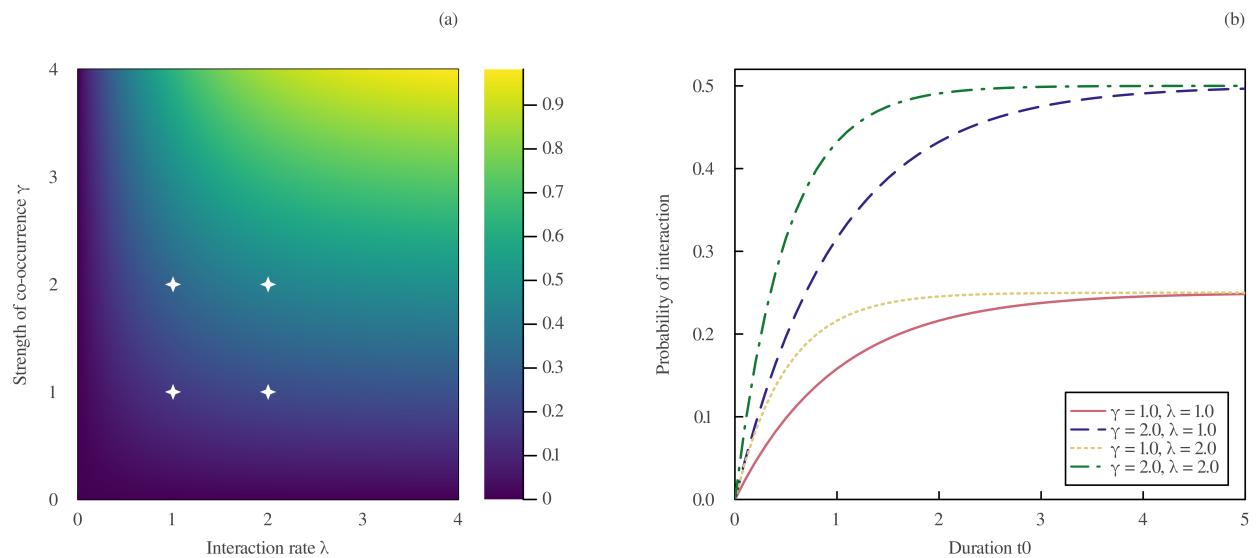


Figure 3: Parameters of the spatiotemporally explicit model of interactions. (a) Probability of local interaction given by the process model (eq. 11) under different values of λ and γ , with $t_0 = 1$. The parameter values used in the right panel are denoted by the white stars. (b) Scaling of the probability of interaction with the duration parameter t_0 in eq. 11, 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.

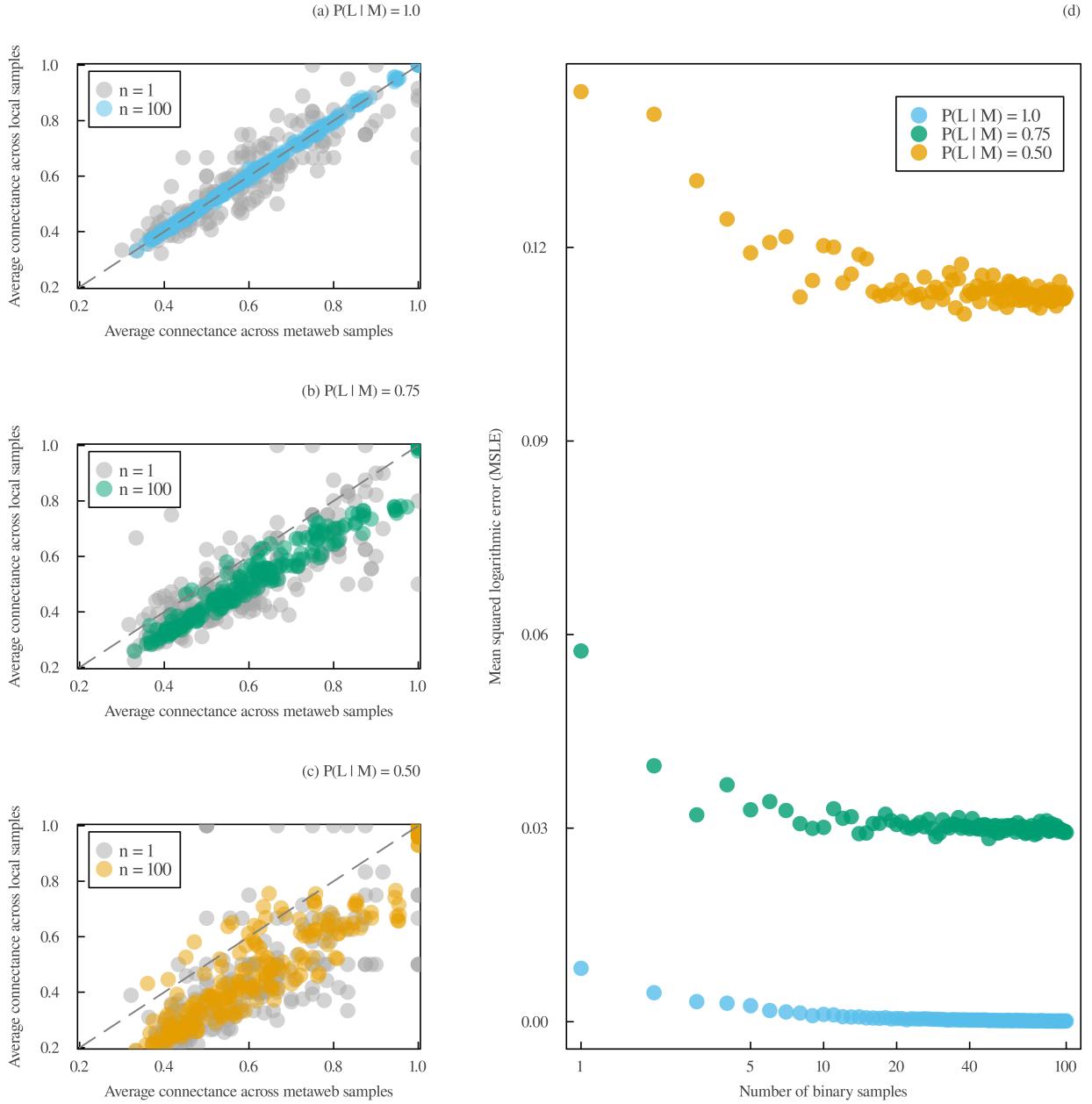


Figure 4: Connectance of sampled binary interaction webs. Comparison between the average connectance of binary interaction network samples obtained from the local and metawebs of probabilistic interactions. Each dot corresponds to a different site. The local probability of interaction between potentially interacting species was set to three different values: (a) $P(L|M) = 1.0$, (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 interaction 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 interaction metaweb, and then propagating this result to all local webs that include the species potentially engaged in the interactions. Local binary interaction webs were generated by independently sampling binary interactions for each local web of probabilistic interactions.