

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 6, 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 abundances. While species' absolute abundance may impact interaction frequencies (Vázquez *et al.*
15 (2007)), encounter probabilities are determined by their relative abundance (Canard *et al.* (2012), Canard *et al.*
16 (2012)). The probability that species meet each other also depends on their biological characteristics, such as
17 the synchronization of their phenology (Olesen *et al.* (2010), Singer & McBride (2012)) and their
18 discoverability (e.g., Broom & Ruxton (2005)). Finally, when species do come into contact, an interaction
19 occurs only if their traits are locally compatible (Poisot *et al.* (2015)), including but not limited to their body
20 phenotypes (Bolnick *et al.* (2011), Stouffer *et al.* (2011), Gravel *et al.* (2013)) and behavioral choices (Pulliam
21 (1974), Choh *et al.* (2012)). Interactions may also be influenced by the presence or prevalence of a third species
22 (e.g., of a more profitable prey species, Golubski & Abrams (2011), Sanders & van Veen (2012)). Documenting
23 the location and timing of interactions becomes even more difficult when accounting for the spatiotemporal
24 variability of ecological interactions (Poisot *et al.* (2012), Poisot *et al.* (2015)). Environmental factors, such as
25 temperature (Angilletta *et al.* (2004)), drought (Woodward *et al.* (2012)), climate change (Gilman *et al.* (2010),
26 Woodward *et al.* (2010), Araujo *et al.* (2011)), and habitat modifications (Tylianakis *et al.* (2007)), contribute
27 to this spatiotemporal variability by impacting species abundance and traits. Even after satisfying all these
28 conditions, there remains a possibility that the interaction does not occur locally, either due to the intricate
29 nature of the system or simply by chance. If it does occur, it might still go unnoticed, particularly if it happens
30 infrequently. In this context, it is unsurprising that our knowledge of ecological interactions remains limited

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

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

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

34 variability of species interactions has led ecologists to expand their representation of ecological networks (also

35 known as ecological webs) to include a probabilistic view of interactions (Poisot *et al.* (2016), Dallas *et al.*

36 (2017), Fu *et al.* (2021)). This different perspective allows us to fill in the Eltonian shortfall (Hortal *et al.*

37 (2015)) by modeling the probability of detecting interactions, which can be an important tool for directing

38 efforts and taking action, especially in places where access and resources for research are scarce. The

39 probabilistic representation of interactions is thus far limited to direct interactions, which are conceptually and

40 mathematically analogous regardless of their biological type (e.g., predation and pollination). This is in contrast

41 with indirect interactions (e.g., interspecific competition), which arise from distinct ecological processes and are

42 often not directly observable (Kéfi *et al.* (2015), Kéfi *et al.* (2016)). Representing direct interactions

43 probabilistically enables us to capture the spatiotemporal variability of the aforementioned ecological processes

44 and the uncertainties associated with their measurement. As opposed to webs of binary deterministic

45 interactions, in which interactions are regarded as either occurring or not, webs of probabilistic interactions,

46 within a Bayesian framework, express our degree of belief (or confidence) regarding the occurrence or

47 observation of interactions. Based on the scale at which they are estimated, probabilistic interactions may

48 reflect our level of confidence in whether interactions will be observed, realized locally, or biologically feasible.

49 As an illustration, we could outline a situation in which there is a 50% certainty that an interaction occurs 50%

50 of the time, or that there is a 50% certainty that it simply occurs. Our level of confidence should be more

51 definitive (approaching either 0 or 1) as we extend our sampling to a broader area and over a longer time period,

52 thereby diminishing the uncertainty of the interactions (but not necessarily the estimation of their variability).

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

54 value of an interaction is restrained to 0 (non-occurring) or 1 (occurring). Yet, for the sake of clarity, we omit

55 binary interactions from our discussion of probabilistic interactions in this contribution. In networks of

56 probabilistic interactions, only forbidden interactions (i.e., interactions prohibited by biological traits or species

57 absence, Jordano *et al.* (2003), Olesen *et al.* (2010)) have a probability value of zero by default, provided that

58 intraspecific trait variability is considered (Gonzalez-Varo & Traveset (2016)).

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

60 realistic portrait of species interactions. However, it is still unknown how the uncertainty of pairwise

interactions propagate to network structure (i.e., community-level properties that drive the functioning, dynamics, and resilience of ecosystems, Proulx *et al.* (2005), McCann (2007), McCann (2011), Rooney & McCann (2012)). The application and development of computational methods in network ecology, often based on a probabilistic representation of interactions, can alleviate (and guide) the sampling efforts required to document species interactions (Strydom *et al.* (2021)). For example, statistical models can be used to estimate the uncertainty of pairwise interactions (Cirtwill *et al.* (2019)) and the probability of missing (false negatives) and spurious (false positives) interactions (Guimerà & Sales-Pardo (2009)). Considering the high rate of false negatives in species interaction data due to the difficulty of witnessing rare interactions (Catchen *et al.* (2023)), these models could inform the identification of priority sampling locations of ecological webs where data collection would yield the most valuable information, thereby reducing errors. Optimization models for sampling locations have mostly found applications in biological systems that are not networks, such as identifying priority sampling sites for disease hotspots (Andrade-Pacheco *et al.* (2020)), but there is substantial promise in applying them to probabilistic ecological interactions. Statistical models can also be used to generate predictions of ecological webs without prior knowledge of pairwise interactions, for instance using body size (Petchey *et al.* (2008), Gravel *et al.* (2013)), phylogeny (Elmasri *et al.* (2020), Strydom *et al.* (2022)), or a combination of niche and neutral processes (Bartomeus *et al.* (2016), Pomeranz *et al.* (2019)) for inference.

Topological null models, which generate networks of probabilistic interactions by preserving chosen characteristics of the adjacency matrix of binary interactions while intentionally omitting others (Bascompte *et al.* (2003), Fortuna & Bascompte (2006)), serve as other examples of common probabilistic interaction models.

Null models can be used to produce underlying distributions of network measures for null hypothesis significance testing. Many measures have been developed to describe the structure (Poisot *et al.* (2016)) and diversity (Ohlmann *et al.* (2019), Godsoe *et al.* (2022)) of probabilistic interaction webs. These models and measures support the use of this approach for the study of a wide range of ecological questions, from making better predictions of species distribution (Cazelles *et al.* (2016)) to forecasting the impact of climate change on ecological webs (Gilman *et al.* (2010)).

The lack of clear guidelines on the use of probabilistic interaction data is worrisome, both for data producers and re-users who generate and manipulate these numbers. This is concerning because sampling strategies and decisions regarding network construction can affect our understanding of network properties (Brimacombe *et al.* (2023)). Besides methodological difficulties that may arise when assessing probabilistic interactions, a precise definition of probabilistic interactions appears to be lacking, making the estimation and use of these data

more difficult. We aim to take a step back by outlining different ways in which probabilistic interactions are defined and used in network ecology. We distinguish two broad categories of probabilistic interaction networks that necessitate distinct approaches when applied to key ecological questions: local webs describing probabilities of realized interactions, and regional webs (metawebs) describing probabilities of potential interactions. We highlight the distinctions in the ecological meaning of these two representations and examine some of their properties and relationships (particularly with space, time, and between each other) through empirical case studies. Moreover, there is currently no metadata standard that could guide the documentation of all types of probabilistic interactions (although see e.g., Salim *et al.* (2022) who discuss data standards for deterministic mutualistic webs). Well-defined metadata for probabilistic interactions would support more adequate manipulation and integration of interaction data from different sources and guard against possible misinterpretations arising from ambiguous definitions of probabilistic interaction networks. These metadata should outline the nature (i.e., local or regional) and type (e.g., predatory or pollination) of the interactions, provide information regarding the taxonomic level, identities, and characteristics (e.g., life stages) of the individuals involved in an interaction, present the mathematical formulation of probabilities, including clearly identified conditional variables (e.g., spatial and temporal scales), and describe the methods and contexts (e.g., location, time, environmental conditions) in which interactions were estimated. Inadequately documented probabilistic interaction data should be used with caution when analyzing ecological webs. Our observations and advice can be applied to most types of ecological networks representing direct interactions, from food webs to host-virus networks. Even though the measurement of interaction probabilities may differ between network types (e.g., estimations often based on frequencies of interactions in plant-pollinator networks), the broad principles underlying our findings remain relevant and applicable across diverse ecological contexts.

Probabilistic representations of interactions

One of the first aspects to take into consideration when estimating or interpreting probabilities of interactions is knowing if they describe potential or realized interactions, as these two types of interactions have distinct conceptual underpinnings and sources of uncertainty. A potential (regional) interaction is defined as the biological capacity of two taxa to interact (i.e., the probability that they interact if they were to encounter each other and given sufficient time) whereas a realized (local) interaction is the occurrence or observation of this interaction in a well-defined space and time (i.e., the probability that they interact locally). For two co-occurring

119 taxa and over infinite time, the probability of local interaction is equivalent to the probability of regional
120 (potential) interaction. Our discussion of local interactions focuses on their occurrence rather than their direct
121 empirical observation, as one of the primary goals of describing them probabilistically is to characterize our
122 uncertainty regarding their actual manifestation. We use the terms *metaweb* (Dunne (2006)) to designate
123 regional webs of potential interactions and *local webs* (Poisot *et al.* (2012)) for those of realized interactions.
124 Metawebs are the network analogs of the species pool, where local webs originate from a subset of both species
125 (nodes) and interactions (edges) of the regional metaweb (Saravia *et al.* (2022)). Without clear documentation,
126 it can be challenging to know if published probabilistic interaction webs describe local or regional interactions
127 (tbl. 1 provides examples of studies employing both types of probabilistic interaction networks), or if so-called
128 probabilities are in reality *interaction scores* (i.e., non-probabilistic quantitative interactions analogous to
129 interaction strengths). When probabilistic regional interactions are used and interpreted incorrectly as local
130 interactions (and conversely), this may generate misleading findings during data analysis. We believe that a
131 better understanding of the differences, similarities, and interconnections between these two probabilistic
132 representations of ecological interactions would alleviate interpretation errors (e.g., when studying network-area
133 relationships) and facilitate a more adequate utilization of interaction data.

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

135 Local and metawebs, like any type of network, are made of nodes and edges that can be represented at different
136 levels of organization and precision. The basic units of ecological networks are individuals that interact with
137 each other (e.g., by predation in food webs, Elton (2001)), forming individual-based networks (Melián *et al.*
138 (2011)). The aggregation of these individuals into more or less homogeneous groups (e.g., populations, species,
139 families, feeding guilds) allows us to represent nodes at broader taxonomic scales, which affects our
140 interpretation of the properties of these systems (Guimarães (2020), Hemprich-Bennett *et al.* (2021)).

141 Edges linking nodes can describe a variety of interaction measures. Ecologists have traditionally represented
142 interactions as binary objects that were considered realized after observing at least one individual from group i
143 interact with at least another individual from group j . In an adjacency matrix of binary interactions B , the
144 presence or absence of an interaction $B_{i \rightarrow j}$ between two taxa can be viewed as the result of a Bernoulli process
145 $B_{i \rightarrow j} \sim \text{Bernoulli}(P(B_{i \rightarrow j}))$, with $P(B_{i \rightarrow j})$ being the probability of interaction that characterizes our limited
146 knowledge of the system and its intrinsic spatiotemporal variability. In networks of probabilistic interactions,
147 $P(B_{i \rightarrow j})$ are edge values, and the only two possible outcomes are the presence ($B_{i \rightarrow j} = 1$) or absence ($B_{i \rightarrow j} = 0$)

148 of an interaction between each pair of nodes. Depending on the type of probabilistic interaction network (local
149 or metaweb), the mathematical formulation and interpretation of stochastic parameters like $P(B_{i \rightarrow j})$ can be
150 linked to environmental and biological factors such as species abundance, traits, area, and time (tbl. 1), for
151 example using logistic regression with a logit link function with continuous explanatory variables. Predicting
152 the number of local webs in which the interaction occurs can be achieved by using a Binomial distribution,
153 assuming a constant probability of interaction and independence between networks (trials). When considering
154 uncertainties around the estimation of $P(B_{i \rightarrow j})$, a Beta distribution can also be used to encompass all possible
155 probability values. In that case, a Beta-Binomial distribution can be used to predict the number of networks in
156 which the interaction occurs. Observing an interaction between two taxa at a given location and time provides
157 important information that can be used to update previous estimates of $P(B_{i \rightarrow j})$, informing us on the biological
158 capacity of both taxa to interact and the environmental conditions that enabled them to interact locally.

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

171 Just like binary interaction networks, the uncertainty and spatiotemporal variability of interaction strengths can
172 be represented probabilistically. However, the need to estimate the probability distribution of all possible values
173 of interaction strengths can make the inference of probabilities more challenging in quantitative webs compared
174 to binary interaction webs, which require only one probability estimate for each interaction. Interaction
175 strengths can follow various probability distributions depending on the measure used. For instance, they can
176 follow a Poisson distribution $W_{i \rightarrow j} \sim \text{Poisson}(\lambda_{i \rightarrow j})$ when predicting frequencies of interactions between pairs
177 of nodes, with $\lambda_{i \rightarrow j}$ being the expected rate at which individuals of taxa i and j interact (e.g., the average

number of prey j consumed by all predators i in a given time period). The Poisson distribution can also be 0-inflated after initially modeling non-interacting taxa (e.g., Boulangeat *et al.* (2012) employ a 0-inflated model to analyze species abundance following the modeling of species presence and absence), which constitute the majority of taxa pairs in most local webs due to their typically high sparseness (Jordano (2016)). Because of the methodological difficulties typically encountered when building deterministic quantitative webs (which are only partially mitigated by models such as Ecopath, Plagányi & Butterworth (2004)), binary interaction webs, which are easier to sample (Jordano (2016)) and predict (Strydom *et al.* (2021)), have been more frequently studied and modeled. Moreover, most published networks of probabilistic interactions and methods describe probabilistic interactions whose outcome is binary (whether interaction probabilities are regarded as constant or variable, e.g. represented by a Beta distribution), which underlines the need for better guidelines regarding the interpretation and manipulation of these types of webs first. For these reasons, our primary focus in this contribution will be on addressing the challenges in interpreting and using interaction probabilities in Bernoulli distributions, in both local and metawebs.

Local webs: communities interacting in space and time

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

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

were either observed or predicted. Even though space and time are continuous variables that should yield probability *densities* of interactions (i.e., relative likelihoods of interactions occurring at infinitesimal locations and instants in time), these definitions enable them to be conceptualised as distinct patches and time segments. Treating space and time as discrete dimensions aligns with the common sampling methods of ecological networks and provides actual probabilities of interactions, which can be obtained by integrating probability densities over space and time. Furthermore, we can quantify both an area A and a duration t , which can be readily used in spatiotemporal analyses of ecological networks. 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 $C_{i,j} \sim \text{Bernoulli}(P(C_{i,j}))$. 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)), potentially taking into account biotic interactions (Staniczenko *et al.* (2017)). 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 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

234 these shelters. Like co-occurrence, Ω can also be modeled probabilistically when the stochasticity or
235 uncertainty of environmental and biological factors is considered. In sum, Ω represents all ecological and
236 biological variables that were taken into consideration when measuring interaction probabilities and is,
237 therefore, a subset of all factors impacting ecological interactions. It is a vector of random variables whose
238 values determine the overall uncertainty of an interaction.

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

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

247 The representation of the local context in which probabilities are estimated and the variables that should be
248 taken into consideration depend on the study system, the objectives of the study, and the resources available to
249 the researchers. In other words, these variables do not systematically need to be accounted for. For example, in
250 Gravel *et al.* (2019), for the purpose of model comparison, some models incorporated environmental variables
251 as conditional factors to estimate interaction probabilities, while others did not. When accounted for, these
252 variables should be clearly described in the documentation of the data (Brimacombe *et al.* (2023)),
253 preferentially in mathematical terms to avoid any confusion in their interpretation and to limit manipulation
254 errors during their re-use. For instance, ecologists should be explicit about their consideration of co-occurrence
255 in their estimation of local interaction probabilities. Indeed, it is important to specify if probability values are
256 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
257 interpretation and analysis of the data. In [tbl. 1](#), we present examples of studies that used these diverse
258 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.

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
$P(L_{i \rightarrow j} X_{il}, X_{jl}, \dots)$	local	realization of the interaction given taxa co-occurrence	spatiotemporal variability	Gravel <i>et al.</i> (2019)
$P(L_{i \rightarrow j} N_{il}, N_{jl}, \dots)$	local	realization of the interaction given taxa abundances	neutral models	Canard <i>et al.</i> (2014)
$P(L_{i \rightarrow j} T_{il}, T_{jl}, \dots)$	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) *
$P(L_{i \rightarrow j} t, \dots)$	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)

Expression	Type	Outcome	Uncertainty	Reference
			sources	
$P(L_{i \rightarrow j} M_{i \rightarrow j}, \dots)$	local	realization of the interaction given that the taxa can biologically interact	spatiotemporal variability	this study
$P(O_{i \rightarrow j} L_{i \rightarrow j}, \dots)$	local	observation of the interaction given that it is realized locally	sampling model	Catchen <i>et al.</i> (2023)

259 Metawebs: regional catalogs of interactions

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

267 Regional interactions are not evaluated for a particular local context. In contrast with local webs of probabilistic
 268 interactions, which represent the stochasticity of interactions occurring in nature, metawebs of probabilistic
 269 interactions measure our degree of belief in the capacity of two taxa to interact (i.e., the probability that their
 270 traits could support an interaction in the right conditions). Consequently, although *neutrally* forbidden
 271 interactions (i.e., improbable interactions between rare species, Canard *et al.* (2012)) tend to have low
 272 probability values in local webs, they may exhibit a higher probability in the metaweb. Potential interactions
 273 describe the probability that there exists at least one combination of phenotypes of taxa i and j that can interact
 274 with each other if they were to encounter. For example, a piscivorous species may have the capacity to interact
 275 with a potential prey species if they share at least one matching combination of body sizes, such as an adult
 276 predator being able to consume juvenile prey. To reduce any biases in our calculation of this probability, it is
 277 crucial to ensure that the set of traits sampled or considered accurately reflects the overall trait distribution in
 278 both taxa. This enhances our confidence in concluding that the taxa cannot interact when no positive
 279 observation has been made and increases our capacity to detect an interaction when it is biologically feasible. In

280 addition to pairwise interactions, the extent of sampling effort also contributes to shaping our assessment of
281 metaweb properties, as sampling a larger number of local webs allows us to capture more regional interactions
282 (McLeod *et al.* (2021)). A probability of potential interaction in a metaweb M can be expressed as

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

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

285 Starting from a selected set of taxa, which are usually distributed within a broad region of interest, metawebs
286 can be built using different data sources, including literature review (e.g., Maiorano2020Tetraeu), aggregated
287 local web data (e.g., Gravel *et al.* (2019), Saravia *et al.* (2022)), and predictive models (e.g., Strydom *et al.*
288 (2022), Shaw *et al.* (2024)). Every pair of taxa that have confidently been observed to interact at least once can
289 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
290 webs of probabilistic interactions, in which interaction events usually remain stochastic (i.e., $P(L_{i \rightarrow j}) < 1$) after
291 empirically observing interactions because of their intrinsic spatiotemporal variability. Similarly, although rare
292 interactions typically have low probabilities in local webs, they can have high probabilities in metawebs when
293 the traits of both taxa are congruent. On the other hand, interactions that were never observed can have low
294 probability values in both local and metawebs, going as low as 0 for forbidden links (i.e., interactions forbidden
295 by biological traits). However, many observations of interactions are in reality false positives because of
296 observation errors due to taxonomic misidentifications and ecological misinterpretations (e.g., due to cryptic
297 species and interactions, Pringle & Hutchinson (2020)). Likewise, forbidden interactions can be false negatives
298 in metawebs, e.g. if they have been assessed for specific phenotypes, locations or time. Employing Bayesian
299 models, whether they are mechanistic or phenomenological, could improve the accuracy of our estimation of
300 interaction probabilities within both local and regional webs (e.g., Bartomeus *et al.* (2016), Cirtwill *et al.*
301 (2019)). This improvement is achieved in metawebs by leveraging prior information regarding the feasibility of
302 interactions along with the empirical data on observed interactions.

303 **Properties of probabilistic interaction webs**

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

313 **Host-parasite network data**

314 We use the collection of tripartite host-parasite networks sampled across Europe, created by Kopelke *et al.*
315 (2017), in the following case studies. This dataset contains well-resolved binary local interactions between
316 willows (52 species), willow-galling sawflies (96 species), and their parasitoids (126 species). Given its
317 replicated webs spanning large spatiotemporal scales, this dataset is well-suited for analyzing a variety of
318 ecological hypotheses and processes. Out of a total of 374 local webs, we retained those containing at least 5
319 species, resulting in a set of 233 georeferenced local webs (networks sampled within areas of 0.1 to 0.3 km²
320 during June and/or July spanning 29 years). We built a metaweb of binary interactions by aggregating all local
321 interactions, which gave us a regional web composed of 274 species and 1080 interactions. In the first two
322 panels of fig. 1, we show how the dissimilarity of interactions between common species (β_{OS}) and the
323 dissimilarity in species composition (β_S) between the metaweb and aggregated local webs (Poisot *et al.* (2012))
324 vary with the number of sampled local webs. This shows that networks of local interactions are highly
325 dissimilar from the metaweb, both in terms of species and interactions, especially when only a limited number
326 of sites has been sampled. Both dissimilarity indices were calculated based on the number of items shared by
327 the two webs (c_{LM}) and the number of items unique to the metaweb (u_M) and to the aggregated local web (u_L).
328 The β_S dissimilarity index uses species (nodes) as items being compared, while the β_{OS} index assesses
329 dissimilarity based on interactions between shared species (Poisot *et al.* (2012)). Both indices were calculated
330 following the β_W index of Whittaker (1960):

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

331

[Figure 1 about here.]

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

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

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

351 We built local webs of probabilistic interactions from the binary ones by using the metawebs of probabilistic
 352 interactions and a constant value of $P(L_{i \rightarrow j}|M_{i \rightarrow j})$ across interactions. Low values of $P(L_{i \rightarrow j}|M_{i \rightarrow j})$ indicate that
 353 feasible interactions rarely occur locally, intermediate values around 50% suggest considerable spatiotemporal
 354 variability, while high values indicate that regional interactions are nearly always realized locally. Following

355 eq. 5, the local probability of interaction between a given pair of taxa consistently remained equal to or below
356 their probability of regional interaction.

357 In the last two panels of fig. 1, we show how the aggregated number of links and connectance (i.e., the
358 proportion of all of the non-forbidden links that are realized) scale with the number of sampled local webs,
359 according to different values of $P(L_{i \rightarrow j} | M_{i \rightarrow j})$. When aggregating local webs of probabilistic interactions, the
360 constancy of the probability of regional interaction across the entire study area means that any rise in the
361 probability of local interaction is solely attributable to an increase in $P(L_{i \rightarrow j} | M_{i \rightarrow j})$.

362 For example, let L_1 and L_2 be two local networks and L_{12} the aggregated web. If $P(L_{1,i \rightarrow j} | M_{i \rightarrow j})$ and
363 $P(L_{2,i \rightarrow j} | M_{i \rightarrow j})$ are the probabilities that two potentially interacting species interact respectively in L_1 and L_2 ,
364 the probability $P(L_{12,i \rightarrow j} | M_{i \rightarrow j})$ that these species interact in the aggregated web L_{12} is obtained by:

$$P(L_{12,i \rightarrow j} | M_{i \rightarrow j}) = 1 - (1 - P(L_{1,i \rightarrow j} | M_{i \rightarrow j})) \times (1 - P(L_{2,i \rightarrow j} | M_{i \rightarrow j})), \quad (6)$$

365 assuming independence between the interactions of two species in different networks. This equation represents
366 the probability that the interaction is realized in either (1) exclusively the local web L_1 , (2) exclusively the local
367 web L_2 or (3) both local webs, given that the two taxa have the biological capacity to interact.

368 By comparing the scaling relationships observed in webs of binary and probabilistic interactions, fig. 1
369 illustrates that high values of $P(L_{i \rightarrow j} | M_{i \rightarrow j})$ lead to systematic overestimations in the number of links and
370 connectance, especially when $P(L_{i \rightarrow j} | M_{i \rightarrow j}) = 1$ (corresponding to the scenario where local probabilities of
371 interactions are equivalent to the probabilities of regional interactions). However, these biases tend to diminish
372 as the number of sampled webs increases.

373 Spatial and temporal scaling of interactions

374 The investigation of network-area relationships and interaction accumulation curves is an important area of
375 research in network ecology. First, network-area relationships elucidate the scaling of network properties (such
376 as modularity and trophic chain lengths) with spatial scale (e.g., Wood *et al.* (2015), Galiana *et al.* (2018)). The
377 variations in network structure across spatial scales may stem from the scaling of species richness (species-area
378 relationships, SARs), the number of interactions (Brose *et al.* (2004)), and many other higher-level properties of
379 the system (e.g., environmental heterogeneity, Thompson & Townsend (2005)) with the sampled area.

380 Additionally, ecological processes unfolding at distinct spatial scales, such as the spatial variability in local
381 community composition resulting from different sequences of extinction and colonization events, can also
382 contribute to this variation (Galiana *et al.* (2018)). Next, interaction accumulation curves describe the scaling of
383 the number of observed interactions with sampling effort (Jordano (2016)). Sampling effort, which may
384 correspond to the duration of the sampling period used to construct the network, can impact connectance
385 (Bersier *et al.* (1999)) and various measures of network structure (Banašek-Richter *et al.* (2004), McLeod *et al.*
386 (2021)). Apart from sampling effort, the temporal scaling of interactions also elucidates how network structure
387 changes with the temporal resolution of the network (Poisot *et al.* (2012)), acknowledging that distinct
388 interactions take place over time, ranging from short-term fluctuations of interactions to long-term trends. As
389 local webs of probabilistic interactions may explicitly account for the spatiotemporal variability of interactions,
390 they offer a distinct approach to investigating the scaling of network structure with space and time, in contrast to
391 webs of binary and quantitative interactions, by making the stochasticity of interactions the focal point of the
392 modeling process.

393 Local and metawebs exhibit distinct relationships with spatial and temporal scales. On one hand, metawebs of
394 probabilistic interactions, being independent of any local context, feature regional interactions that do not scale
395 with space and time. This is because regional interactions depend solely on the biological capacity of two taxa
396 to interact, regardless of their co-occurrence and specific environmental conditions. However, probabilities of
397 potential interactions may change (tending to become more definitive) upon updating previous estimates with
398 increased sampling effort, even though they do not vary in a specific direction with the spatial and temporal
399 extent (boundaries) of the network. The probability of two taxa potentially interacting should theoretically be
400 the same in all metawebs in which they are present, provided that the data and methods used for estimation are
401 consistent. For example, if a smaller metaweb M_1 is derived from a larger metaweb M_0 by selecting the subset
402 of taxa present in the region described by M_1 and retaining all their interactions, their probabilities of
403 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
404 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
405 larger web, even though pairwise probabilities remain identical.

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

410 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
411 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,
412 interaction probabilities could be higher in the smaller area, contingent upon their environmental and biological
413 conditions. Likewise, interaction probabilities are expected to be lower in webs with shorter durations when
414 time intervals are nested. In fig. 2, we show how the expected number of local host-parasite interactions scales
415 with area, represented as an expanding latitudinal window, in comparison with regional interactions. Even
416 though we employed local probabilities of interactions equal to regional interactions for the purpose of
417 comparison (i.e., using $P(L|M) = 1$ here), we notice that the total number of regional interactions scales more
418 rapidly than local interactions. This is because numerous regional interactions involve species that never
419 co-occur, and as a result, these interactions are not captured in local webs.

420 [Figure 2 about here.]

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

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

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

438 co-occurrence (Cazelles *et al.* (2016)):

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

439 When $\gamma > 1$, it signifies a positive association in the geographic distributions of both taxa, indicating that the
440 presence of one taxon enhances the probability of occurrence of the other. Attractions may be the result of
441 positive interactions (e.g., mutualism) or positive/negative interactions (e.g., antagonism between predators and
442 prey, Cazelles *et al.* (2016)). In empirical webs, $\gamma > 1$ holds true for the majority of species pairs (Catchen *et*
443 *al.* (2023)). In contrast, repulsions ($\gamma < 1$) may be caused by strong interspecific competition (Cazelles *et al.*
444 (2016)). We model the co-occurrence of both taxa as the outcome of a Bernoulli trial

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

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

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

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

454 The occurrence of an interaction between i and j can be modeled as a Bernoulli trial with a probability of
455 $P(L_{i \rightarrow j})$. Consequently, a Bayesian inference model can be built based on the preceding equations to estimate
456 the value of the λ and γ parameters and generate novel interaction data:

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

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

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

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

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

462 [Figure 3 about here.]

463 Taxonomic scaling of interactions

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

473 There are no inherent differences in the taxonomic scaling between local and metawebs, as only the nodes are
 474 defined taxonomically. In other words, the probability values of edges in both local (eq. 2) and metawebs (eq. 3)
 475 are not conditional on any taxonomic level. The taxonomic scale is tied to the definition of the event itself (i.e.,
 476 the interaction between two taxa, defined at the desired taxonomical scale), not to the variables on which

477 interaction probabilities are conditional. In both types of webs, transitioning to a broader level of organization
 478 (e.g., from a species-level web S to a genus-level web G) can be accomplished directly by using probabilities
 479 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
 480 can compute the probability that at least one species from genus g_1 interacts with at least one species from
 481 genus g_2 (i.e., that there is a non-zero number of species-level interactions or, equivalently, that the genus-level
 482 interaction occurs) as follows:

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

483 where g_{1i} and g_{2j} are the species of the corresponding genus and assuming independence between species-level
 484 interactions. If it is known that at least two of these species interact (i.e., $P(S_{g_1 i \rightarrow g_2 j}) = 1$ for at least one pair of
 485 (g_{1i}, g_{2j})), it implies a probability of genus interaction equal to 1. Canard *et al.* (2012) built a species-based
 486 network following a similar approach, by using simulated interactions between individuals derived from a
 487 neutral model (i.e., a model that assumed ecological equivalence among individuals). In contrast, a more
 488 sophisticated approach is necessary when transitioning from a broader to a finer level of organization. This is
 489 because knowledge of an interaction between two genera does not guarantee that all possible pairwise
 490 combinations of their species will also interact. One possible method is to build a finer-scale network by
 491 generating probabilities of interactions through random sampling from a beta distribution, parameterized by the
 492 broader-scale network.

493 Ideally, our biological interpretation of probabilistic interactions should remain consistent across a network
 494 even if it incorporates heterogeneous levels of organization, such as a network whose nodes represent both
 495 species and trophic species (groups of species sharing similar predators and prey). This is common in
 496 ecological webs where taxonomic resolution is typically low (Hemprich-Bennett *et al.* (2021), Vázquez§ *et al.*
 497 (2022)). Interaction probabilities at broader taxonomic scales should be based on probabilities of interactions
 498 between individuals, for both local and metawebs. For instance, in local individual-based food webs, the
 499 probability that two individuals interact reflects our degree of belief that one individual will consume the other.
 500 Similarly, in local species-based food webs, the probability that two species interact represents our degree of
 501 belief that at least one individual from the predator species consumes at least another individual from the prey
 502 species. Furthermore, in local clade-based food webs, the probability that two clades interact represents our
 503 degree of belief that at least two species from these clades interact with each other or, equivalently, that at least

504 two individuals from these clades interact with each other. Fundamentally, the taxonomic scaling of interactions
505 involves aggregating interactions between individuals into larger groups, which may exhibit varying degrees of
506 homogeneity based on the organisms and the study system. In that regard, taxonomic scaling is analogous to the
507 spatial and temporal scaling of interactions, as they all represent different ways to aggregate individuals into
508 broader groups (either spatially, temporally, or taxonomically).

509 Sampling for binary interaction webs

510 The prediction of binary interactions through Bernoulli trials is an important application of probabilistic
511 interaction webs. This approach proves beneficial for analyzing the structural characteristics of probabilistic
512 interaction webs, particularly in the absence of specific analytical measures. By performing independent
513 Bernoulli trials for each interaction, a network of binary interactions may be generated. A probability
514 distribution of network properties can be obtained by measuring network structure across multiple randomly
515 generated networks (Poisot *et al.* (2016)). This method enables the representation of the variability of network
516 structure, albeit with possible biases when connectance is low (Poisot & Gravel (2014), Chagnon (2015)).

517 Employing this strategy to generate binary interaction networks under a null model facilitates null hypothesis
518 significance testing, wherein the observed measure is compared against the simulated distribution (e.g.,

519 Bascompte *et al.* (2003)). Additionally, the random generation of binary interaction networks, from a
520 probabilistic interaction web that accounts for the spatiotemporal variability of interactions, may effectively
521 capture network structure across space and time. This facilitates the investigation of ecological hypotheses
522 about interactions at broad spatial and temporal scales. Yet, inferring network structure through sampling for
523 binary interactions assumes independence among interactions, which might not accurately represent reality.

524 Covariation among interactions could exist even without conditioning interactions on others, e.g. an interaction
525 between two taxa could be more probable when another interaction occurs. The consequences of this
526 assumption of independence have yet to be empirically examined.

527 There are at least two distinct approaches to sample binary interaction networks from probabilistic interaction
528 webs across space, for example, when attempting to predict a binary interaction network for each of a number of
529 locations within a given region. Both approaches assume independence between interactions. The first
530 approach involves performing a singular Bernoulli trial for each pair of taxa based on their regional probability
531 of interaction. In employing this approach, every pair of taxa predicted to interact in the metaweb of binary
532 interactions will be treated as interacting in all local webs where they co-occur. This will result in local pairwise

interactions without spatial variation. The second approach is to independently sample each of the local webs of probabilistic interactions, which can be achieved by first generating distinct probabilistic interaction networks for each location. These local webs of probabilistic interactions may vary in taxa composition and interaction probabilities. Subsequently, binary interaction networks can be independently sampled for each location, introducing spatial variation in binary interactions.

In fig. 4, we compare the average connectance of binary interaction networks resulting from these two sampling 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

559 cell size, preventing systematic biases in predictions.

560 Prediction of local webs from metawebs

561 Metawebs serve as a valuable source of ecological information for predicting local webs across time and space.

562 Local webs of binary interactions can be reconstructed by selecting a subset of taxa and interactions from the

563 metaweb (Dunne (2006)). This implies that metawebs consistently contain more interactions than their

564 corresponding local webs. Determining the list of taxa to select can be achieved empirically or through methods

565 like range maps or species distribution models. As the species composition of a community is arguably less

566 difficult to sample or predict than its interactions, the primary challenge lies in deciding which interactions to

567 select from the metaweb, a task that may necessitate advanced statistical models and ecological expertise.

568 Inferring the structure of local webs from the metaweb before predicting specific local pairwise interactions

569 could hold promise (Strydom *et al.* (2021)), considering that the structure of local webs is constrained by the

570 metaweb (Saravia *et al.* (2022)).

571 Building local webs of probabilistic interactions from a metaweb of probabilistic interactions involves a

572 reduction in the value of pairwise interaction probabilities. This decrease is attributed to the prerequisite that

573 two taxa must initially possess the capacity to interact before engaging in local interactions (eq. 5). Therefore,

574 inferring local webs from their metaweb while maintaining identical interaction probability values would

575 introduce systematic biases into the predictions. In such cases, these networks would essentially represent

576 smaller-scale metawebs of potential interactions, possibly leading to misinterpretations by being perceived as

577 local interactions. As proposed by McLeod *et al.* (2021), although metawebs do not capture the spatiotemporal

578 variability of interactions, they establish an upper limit for local interactions. In other words, the probability of

579 two taxa interacting at a specific location and time is consistently lower or equal to the probability of their

580 regional interaction:

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

581 Moreover, the probability that two taxa possess the biological capacity to interact must exceed the probability of

582 them interacting at any location and time because they may never co-occur or encounter locally. Specifically,

583 the cumulative probability of local interactions across all spatial, temporal, and environmental conditions must

584 be less than the probability of regional interaction, i.e.

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

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

590 Conclusion

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

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

622 Acknowledgment

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

630 References

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

- 636 Angilletta, M.J., Jr., Steury, T.D. & Sears, M.W. (2004). Temperature, Growth Rate, and Body Size in
637 Ectotherms: Fitting Pieces of a Life-History Puzzle¹. *Integrative and Comparative Biology*, 44, 498–509.
- 638 Araujo, M.B., Rozenfeld, A., Rahbek, C. & Marquet, P.A. (2011). Using species co-occurrence networks to
639 assess the impacts of climate change. *Ecography*, 34, 897–908.
- 640 Banašek-Richter, C., Cattin, M.-F. & Bersier, L.-F. (2004). Sampling effects and the robustness of quantitative
641 and qualitative food-web descriptors. *Journal of Theoretical Biology*, 226, 23–32.
- 642 Bartomeus, I., Gravel, D., Tylianakis, J.M., Aizen, M.A., Dickie, I.A. & Bernard-Verdier, M. (2016). A
643 common framework for identifying linkage rules across different types of interactions. *Functional Ecology*,
644 30, 1894–1903.
- 645 Bascompte, J., Jordano, P., Melián, C.J. & Olesen, J.M. (2003). The nested assembly of plant–animal
646 mutualistic networks. *Proceedings of the National Academy of Sciences*, 100, 9383–9387.
- 647 Berlow, E.L., Neutel, A.-M., Cohen, J.E., De Ruiter, P.C., Ebenman, B., Emmerson, M., et al. (2004).
648 Interaction strengths in food webs: Issues and opportunities. *Journal of Animal Ecology*, 73, 585–598.
- 649 Bersier, L.-F., Dixon, P. & Sugihara, G. (1999). Scale-Invariant or Scale-Dependent Behavior of the Link
650 Density Property in Food Webs: A Matter of Sampling Effort? *The American Naturalist*, 153, 676–682.
- 651 Blanchet, F.G., Cazelles, K. & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions.
652 *Ecology Letters*, 23, 1050–1063.
- 653 Bohan, D.A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A.J. & Woodward, G. (2017).
654 Next-Generation Global Biomonitoring: Large-scale, Automated Reconstruction of Ecological Networks.
655 *Trends in Ecology & Evolution*, 32, 477–487.
- 656 Bolnick, D.I., Amarasekare, P., Araújo, M.S., Bürger, R., Levine, J.M., Novak, M., et al. (2011). Why
657 intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution*, 26, 183–192.
- 658 Borrett, S.R. & Scharler, U.M. (2019). Walk partitions of flow in Ecological Network Analysis: Review and
659 synthesis of methods and indicators. *Ecological Indicators*, 106, 105451.
- 660 Boulangeat, I., Gravel, D. & Thuiller, W. (2012). Accounting for dispersal and biotic interactions to disentangle
661 the drivers of species distributions and their abundances. *Ecology Letters*, 15, 584–593.
- 662 Brimacombe, C., Bodner, K., Michalska-Smith, M., Poisot, T. & Fortin, M.-J. (2023). Shortcomings of reusing
663 species interaction networks created by different sets of researchers. *PLOS Biology*, 21, e3002068.

- 664 Broom, M. & Ruxton, G.D. (2005). You can run—or you can hide: Optimal strategies for cryptic prey against
665 pursuit predators. *Behavioral Ecology*, 16, 534–540.
- 666 Brose, U., Ostling, A., Harrison, K. & Martinez, N.D. (2004). Unified spatial scaling of species and their
667 trophic interactions. *Nature*, 428, 167–171.
- 668 Canard, E.F., Mouquet, N., Mouillot, D., Stanko, M., Miklisova, D. & Gravel, D. (2014). Empirical Evaluation
669 of Neutral Interactions in Host-Parasite Networks. *The American Naturalist*, 183, 468–479.
- 670 Canard, E., Mouquet, N., Marescot, L., Gaston, K.J., Gravel, D. & Mouillot, D. (2012). Emergence of
671 Structural Patterns in Neutral Trophic Networks. *PLOS ONE*, 7, e38295.
- 672 Catchen, M.D., Poisot, T., Pollock, L.J. & Gonzalez, A. (2023). The missing link: Discerning true from false
673 negatives when sampling species interaction networks.
- 674 Cazelles, K., Araujo, M.B., Mouquet, N. & Gravel, D. (2016). A theory for species co-occurrence in interaction
675 networks. *Theoretical Ecology*, 9, 39–48.
- 676 Chagnon, P.-L. (2015). Characterizing topology of ecological networks along gradients: The limits of metrics'
677 standardization. *Ecological Complexity*, 22, 36–39.
- 678 Choh, Y., Ignacio, M., Sabelis, M.W. & Janssen, A. (2012). Predator-prey role reversals, juvenile experience
679 and adult antipredator behaviour. *Scientific Reports*, 2, 728.
- 680 Cirtwill, A.R., Eklöf, A., Roslin, T., Wootton, K. & Gravel, D. (2019). A quantitative framework for
681 investigating the reliability of empirical network construction. *Methods in Ecology and Evolution*, 10,
682 902–911.
- 683 Dallas, T., Park, A.W. & Drake, J.M. (2017). Predicting cryptic links in host-parasite networks. *PLOS
684 Computational Biology*, 13, e1005557.
- 685 Dunne, J.A. (2006). The Network Structure of Food Webs. In: *Ecological networks: Linking structure and
686 dynamics* (eds. Dunne, J.A. & Pascual, M.). Oxford University Press, pp. 27–86.
- 687 Elmasri, M., Farrell, M.J., Davies, T.J. & Stephens, D.A. (2020). A hierarchical Bayesian model for predicting
688 ecological interactions using scaled evolutionary relationships. *The Annals of Applied Statistics*, 14,
689 221–240.
- 690 Elton, C.S. (2001). *Animal Ecology*. University of Chicago Press, Chicago, IL.

- 691 Emmerson, M.C. & Raffaelli, D. (2004). Predator–prey body size, interaction strength and the stability of a real
692 food web. *Journal of Animal Ecology*, 73, 399–409.
- 693 Fortuna, M.A. & Bascompte, J. (2006). Habitat loss and the structure of plant–animal mutualistic networks.
694 *Ecology Letters*, 9, 281–286.
- 695 Fu, X., Seo, E., Clarke, J. & Hutchinson, R.A. (2021). Link Prediction Under Imperfect Detection:
696 Collaborative Filtering for Ecological Networks. *IEEE Transactions on Knowledge and Data Engineering*,
697 33, 3117–3128.
- 698 Galiana, N., Lurgi, M., Claramunt-López, B., Fortin, M.-J., Leroux, S., Cazelles, K., et al. (2018). The spatial
699 scaling of species interaction networks. *Nature Ecology & Evolution*, 2, 782–790.
- 700 Gilman, S.E., Urban, M.C., Tewksbury, J., Gilchrist, G.W. & Holt, R.D. (2010). A framework for community
701 interactions under climate change. *Trends in Ecology & Evolution*, 25, 325–331.
- 702 Godsoe, W., Murray, R. & Iritani, R. (2022). Species interactions and diversity: A unified framework using Hill
703 numbers. *Oikos*, n/a, e09282.
- 704 Golubski, A.J. & Abrams, P.A. (2011). Modifying modifiers: What happens when interspecific interactions
705 interact? *Journal of Animal Ecology*, 80, 1097–1108.
- 706 Gomez, J.M., Verdu, M. & Perfectti, F. (2010). Ecological interactions are evolutionarily conserved across the
707 entire tree of life. *Nature*, 465, 918–U6.
- 708 Gonzalez, A. & Londoño, M.C. (2022). Monitor biodiversity for action. *Science*, 378, 1147–1147.
- 709 Gonzalez-Varo, J.P. & Traveset, A. (2016). The Labile Limits of Forbidden Interactions. *Trends in Ecology &*
710 *Evolution*, 31, 700–710.
- 711 Gravel, D., Albouy, C. & Thuiller, W. (2016). The meaning of functional trait composition of food webs for
712 ecosystem functioning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371,
713 20150268.
- 714 Gravel, D., Baiser, B., Dunne, J.A., Kopelke, J.-P., Martinez, N.D., Nyman, T., et al. (2019). Bringing Elton
715 and Grinnell together: A quantitative framework to represent the biogeography of ecological interaction
716 networks. *Ecography*, 42, 401–415.
- 717 Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. (2013). Inferring food web structure from
718 predator–prey body size relationships. *Methods in Ecology and Evolution*, 4, 1083–1090.

- 719 Guimarães, P.R. (2020). The Structure of Ecological Networks Across Levels of Organization. *Annual Review*
720 *of Ecology, Evolution, and Systematics*, 51, 433–460.
- 721 Guimerà, R. & Sales-Pardo, M. (2009). Missing and spurious interactions and the reconstruction of complex
722 networks. *Proceedings of the National Academy of Sciences*, 106, 22073–22078.
- 723 Hemprich-Bennett, D.R., Oliveira, H.F.M., Le Comber, S.C., Rossiter, S.J. & Clare, E.L. (2021). Assessing the
724 impact of taxon resolution on network structure. *Ecology*, 102, e03256.
- 725 Herrera, C.M. (1989). Pollinator abundance, morphology, and flower visitation rate: Analysis of the “quantity”
726 component in a plant-pollinator system. *Oecologia*, 80, 241–248.
- 727 Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M. & Ladle, R.J. (2015). Seven Shortfalls
728 that Beset Large-Scale Knowledge of Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*,
729 46, 523–549.
- 730 Jordano, P. (1987). Patterns of Mutualistic Interactions in Pollination and Seed Dispersal: Connectance,
731 Dependence Asymmetries, and Coevolution. *The American Naturalist*, 129, 657–677.
- 732 Jordano, P. (2016). Sampling networks of ecological interactions. *Functional Ecology*, 30, 1883–1893.
- 733 Jordano, P., Bascompte, J. & Olesen, J.M. (2003). Invariant properties in coevolutionary networks of
734 plant-animal interactions. *Ecology Letters*, 6, 69–81.
- 735 Kéfi, S., Berlow, E.L., Wieters, E.A., Joppa, L.N., Wood, S.A., Brose, U., et al. (2015). Network structure
736 beyond food webs: Mapping non-trophic and trophic interactions on Chilean rocky shores. *Ecology*, 96,
737 291–303.
- 738 Kéfi, S., Berlow, E.L., Wieters, E.A., Navarrete, S.A., Petchey, O.L., Wood, S.A., et al. (2012). More than a
739 meal... integrating non-feeding interactions into food webs. *Ecology Letters*, 15, 291–300.
- 740 Kéfi, S., Miele, V., Wieters, E.A., Navarrete, S.A. & Berlow, E.L. (2016). How Structured Is the Entangled
741 Bank? The Surprisingly Simple Organization of Multiplex Ecological Networks Leads to Increased
742 Persistence and Resilience. *PLOS Biology*, 14, e1002527.
- 743 Klecka, J. & Boukal, D.S. (2014). The effect of habitat structure on prey mortality depends on predator and
744 prey microhabitat use. *Oecologia*, 176, 183–191.
- 745 Kopelke, J.-P., Nyman, T., Cazelles, K., Gravel, D., Vissault, S. & Roslin, T. (2017). Food-web structure of
746 willow-galling sawflies and their natural enemies across Europe. *Ecology*, 98, 1730–1730.

- 747 Kulkarni, M.A., Kweka, E., Nyale, E., Lyatuu, E., Mosha, F.W., Chandramohan, D., *et al.* (2006).
748 Entomological Evaluation of Malaria Vectors at Different Altitudes in Hai District, Northeastern Tanzania.
749 *Journal of Medical Entomology*, 43, 580–588.
- 750 McCann, K. (2007). Protecting biostructure. *Nature*, 446, 29–29.
- 751 McCann, K.S. (2011). Food webs (MPB-50). In: *Food Webs (MPB-50)*. Princeton University Press.
- 752 McLeod, A., Leroux, S.J., Gravel, D., Chu, C., Cirtwill, A.R., Fortin, M.-J., *et al.* (2021). Sampling and
753 asymptotic network properties of spatial multi-trophic networks. *Oikos*, n/a.
- 754 McLeod, A.M., Leroux, S.J. & Chu, C. (2020). Effects of species traits, motif profiles, and environment on
755 spatial variation in multi-trophic antagonistic networks. *Ecosphere*, 11, e03018.
- 756 Melián, C.J., Vilas, C., Baldó, F., González-Ortegón, E., Drake, P. & Williams, R.J. (2011). Eco-evolutionary
757 Dynamics of Individual-Based Food Webs. In: *Advances in Ecological Research*, The Role of Body Size in
758 Multispecies Systems (ed. Belgrano, A.). Academic Press, pp. 225–268.
- 759 Ohlmann, M., Miele, V., Dray, S., Chalmandrier, L., O'Connor, L. & Thuiller, W. (2019). Diversity indices for
760 ecological networks: A unifying framework using Hill numbers. *Ecology Letters*, 22, 737–747.
- 761 Olesen, J.M., Bascompte, J., Dupont, Y.L., Elberling, H., Rasmussen, C. & Jordano, P. (2010). Missing and
762 forbidden links in mutualistic networks. *Proceedings of the Royal Society B: Biological Sciences*, 278,
763 725–732.
- 764 Pascual, M., Dunne, J.A. & Dunne, J.A. (2006). *Ecological Networks: Linking Structure to Dynamics in Food
765 Webs*. Oxford University Press, USA.
- 766 Petchey, O.L., Beckerman, A.P., Riede, J.O. & Warren, P.H. (2008). Size, foraging, and food web structure.
767 *Proceedings of the National Academy of Sciences*, 105, 4191–4196.
- 768 Pilosof, S., Porter, M.A., Pascual, M. & Kéfi, S. (2017). The multilayer nature of ecological networks. *Nature
769 Ecology & Evolution*, 1, 1–9.
- 770 Plagányi, É.E. & Butterworth, D.S. (2004). A critical look at the potential of Ecopath with ecosim to assist in
771 practical fisheries management. *African Journal of Marine Science*, 26, 261–287.
- 772 Poisot, T., Canard, E., Mouillot, D., Mouquet, N. & Gravel, D. (2012). The dissimilarity of species interaction
773 networks. *Ecology Letters*, 15, 1353–1361.

- 774 Poisot, T., Cirtwill, A.R., Cazelles, K., Gravel, D., Fortin, M.-J. & Stouffer, D.B. (2016). [The structure of probabilistic networks](#). *Methods in Ecology and Evolution*, 7, 303–312.
- 775
- 776 Poisot, T. & Gravel, D. (2014). [When is an ecological network complex? Connectance drives degree distribution and emerging network properties](#). *PeerJ*, 2, e251.
- 777
- 778 Poisot, T., Stouffer, D.B. & Gravel, D. (2015). [Beyond species: Why ecological interaction networks vary through space and time](#). *Oikos*, 124, 243–251.
- 779
- 780 Polis, G.A. (1991). [Complex Trophic Interactions in Deserts: An Empirical Critique of Food-Web Theory](#). *The American Naturalist*, 138, 123–155.
- 781
- 782 Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O’Hara, R.B., Parris, K.M., *et al.* (2014). [Understanding co-occurrence by modelling species simultaneously with a Joint Species DistributionModel \(JSDM\)](#).
- 783
- 784 *Methods in Ecology and Evolution*, 5, 397–406.
- 785
- 786 Pomeranz, J.P.F., Thompson, R.M., Poisot, T. & Harding, J.S. (2019). [Inferring predator–prey interactions in food webs](#). *Methods in Ecology and Evolution*, 10, 356–367.
- 787
- 788 Pringle, R.M. & Hutchinson, M.C. (2020). [Resolving Food-Web Structure](#). *Annual Review of Ecology, Evolution, and Systematics*, 51, 55–80.
- 789
- 790 Proulx, S.R., Promislow, D.E.L. & Phillips, P.C. (2005). [Network thinking in ecology and evolution](#). *Trends in Ecology & Evolution*, SPECIAL ISSUE: BUMPER BOOK REVIEW, 20, 345–353.
- 791
- 792 Pulliam, H.R. (1974). [On the Theory of Optimal Diets](#). *The American Naturalist*, 108, 59–74.
- 793
- 794 Rooney, N. & McCann, K.S. (2012). [Integrating food web diversity, structure and stability](#). *Trends in Ecology & Evolution*, 27, 40–46.
- 795
- 796 Salim, J.A., Saraiva, A.M., Zermoglio, P.F., Agostini, K., Wolowski, M., Drucker, D.P., *et al.* (2022). [Data standardization of plant–pollinator interactions](#). *GigaScience*, 11, giac043.
- 797
- 798 Sanders, D. & van Veen, F.J.F. (2012). [Indirect commensalism promotes persistence of secondary consumer species](#). *Biology Letters*, 8, 960–963.
- 799 Saravia, L.A., Marina, T.I., Kristensen, N.P., De Troch, M. & Momo, F.R. (2022). [Ecological network assembly: How the regional metaweb influences local food webs](#). *Journal of Animal Ecology*, 91, 630–642.

- 800 Schmeller, D.S., Julliard, R., Bellingham, P.J., Böhm, M., Brummitt, N., Chiarucci, A., *et al.* (2015). Towards a
801 global terrestrial species monitoring program. *Journal for Nature Conservation*, 25, 51–57.
- 802 Shaw, J.O., Dunhill, A.M., Beckerman, A.P., Dunne, J.A. & Hull, P.M. (2024). A framework for reconstructing
803 ancient food webs using functional trait data.
- 804 Singer, M.C. & McBride, C.S. (2012). Geographic mosaics of species' association: A definition and an
805 example driven by plant–insect phenological synchrony. *Ecology*, 93, 2658–2673.
- 806 Staniczenko, P.P.A., Sivasubramaniam, P., Suttle, K.B. & Pearson, R.G. (2017). Linking macroecology and
807 community ecology: Refining predictions of species distributions using biotic interaction networks.
808 *Ecology Letters*, 20, 693–707.
- 809 Stouffer, D.B., Rezende, E.L. & Amaral, L.A.N. (2011). The role of body mass in diet contiguity and food-web
810 structure. *Journal of Animal Ecology*, 80, 632–639.
- 811 Strydom, T., Bouskila, S., Banville, F., Barros, C., Caron, D., Farrell, M.J., *et al.* (2022). Food web
812 reconstruction through phylogenetic transfer of low-rank network representation. *Methods in Ecology and
813 Evolution*, 13.
- 814 Strydom, T., Catchen, M.D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., *et al.* (2021). A
815 roadmap towards predicting species interaction networks (across space and time). *Philosophical
816 Transactions of the Royal Society B-Biological Sciences*, 376, 20210063.
- 817 Thompson, R.M. & Townsend, C.R. (2005). Food-Web Topology Varies with Spatial Scale in a Patchy
818 Environment. *Ecology*, 86, 1916–1925.
- 819 Tylianakis, J.M., Tscharntke, T. & Lewis, O.T. (2007). Habitat modification alters the structure of tropical
820 host–parasitoid food webs. *Nature*, 445, 202–205.
- 821 Vázquez, D.P., Melián, C.J., Williams, N.M., Blüthgen, N., Krasnov, B.R. & Poulin, R. (2007). Species
822 abundance and asymmetric interaction strength in ecological networks. *Oikos*, 116, 1120–1127.
- 823 Vázquez§, D.P., Peralta§, G., Cagnolo, L., Santos, M. & Igual, §.E. autores contribuyeron por. (2022).
824 Ecological interaction networks. What we know, what we don't, and why it matters. *Ecología Austral*, 32,
825 670–697.
- 826 Weinstein, B.G. & Graham, C.H. (2017a). On comparing traits and abundance for predicting species
827 interactions with imperfect detection. *Food Webs*, 11, 17–25.

- 828 Weinstein, B.G. & Graham, C.H. (2017b). Persistent bill and corolla matching despite shifting temporal
829 resources in tropical hummingbird-plant interactions. *Ecology Letters*, 20, 326–335.
- 830 Whittaker, R.H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological
831 Monographs*, 30, 279–338.
- 832 Wood, S.A., Russell, R., Hanson, D., Williams, R.J. & Dunne, J.A. (2015). Effects of spatial scale of sampling
833 on food web structure. *Ecology and Evolution*, 5, 3769–3782.
- 834 Woodward, G., Benstead, J.P., Beveridge, O.S., Blanchard, J., Brey, T., Brown, L.E., et al. (2010). Chapter 2 -
835 Ecological Networks in a Changing Climate. In: *Advances in Ecological Research*, Ecological Networks
836 (ed. Woodward, G.). Academic Press, pp. 71–138.
- 837 Woodward, G., Brown, L.E., Edwards, F.K., Hudson, L.N., Milner, A.M., Reuman, D.C., et al. (2012). Climate
838 change impacts in multispecies systems: Drought alters food web size structure in a field experiment.
839 *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 2990–2997.

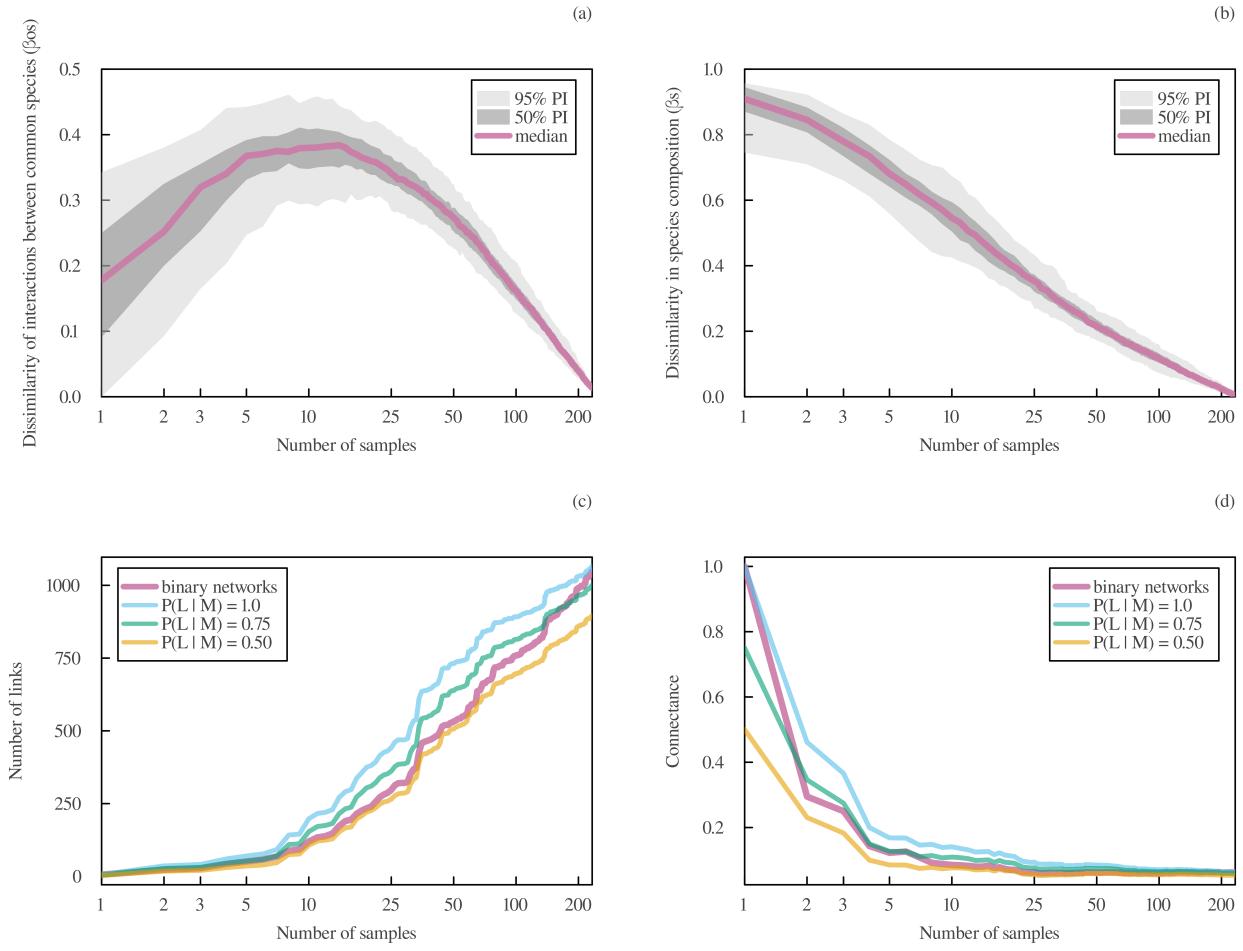


Figure 1: Network accumulation curves. (a) Dissimilarity of interactions between common species and (b) dissimilarity in species composition between aggregated local webs and the metaweb of binary host-parasite interactions. Aggregated local webs were obtained by sequentially and randomly selecting a 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 $P(L_{i \rightarrow j}|M_{i \rightarrow j})$ (the local probability of interaction among potentially interacting species, indices omitted in the figure for simplicity) was used for all local 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).

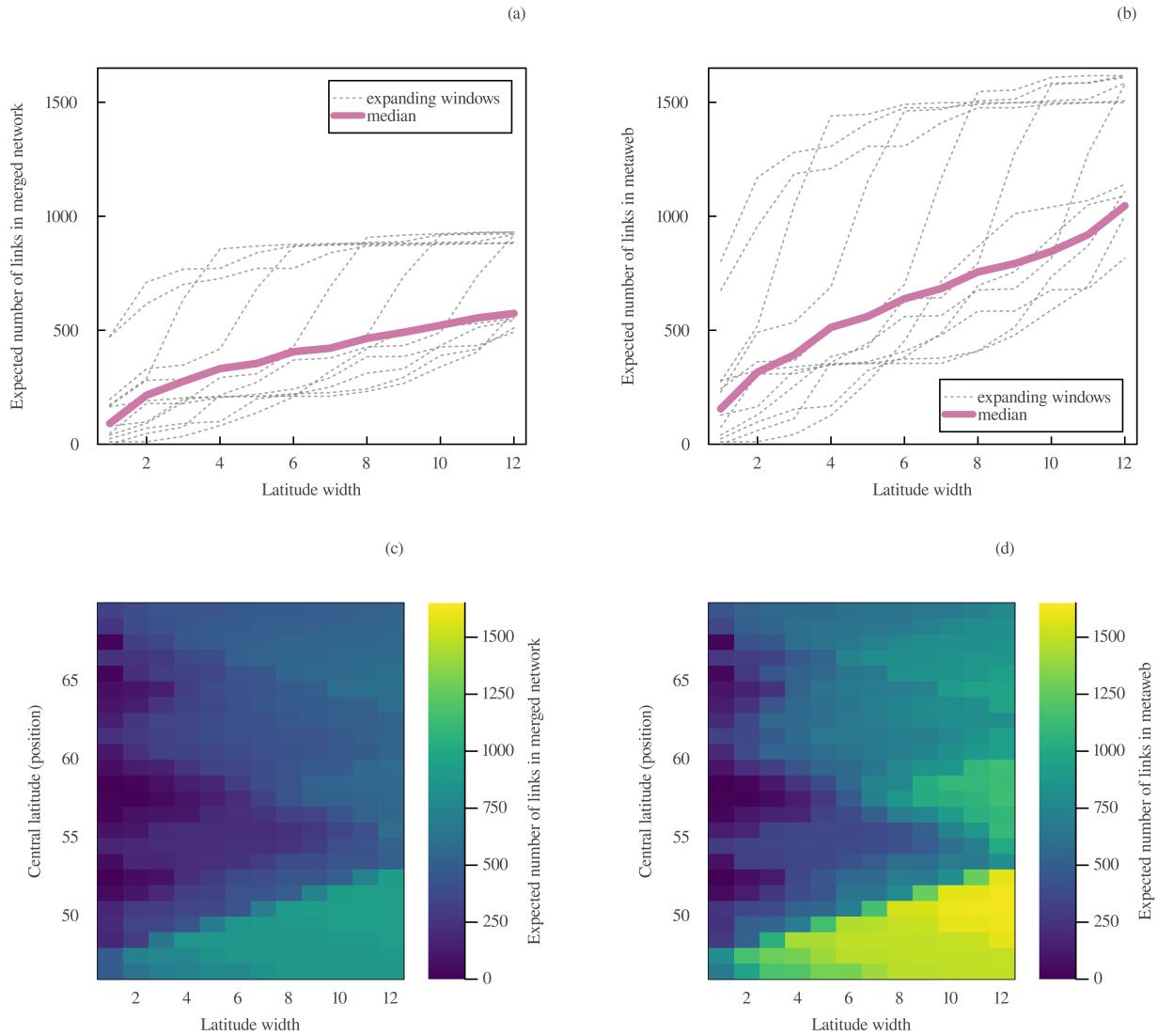


Figure 2: Spatial scaling of interactions. Expected number of host-parasite interactions in a network aggregating all (a) local and (b) regional probabilistic interactions within a latitudinal window of a given length. Every dashed curve corresponds to a different expanding window characterized by a 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.

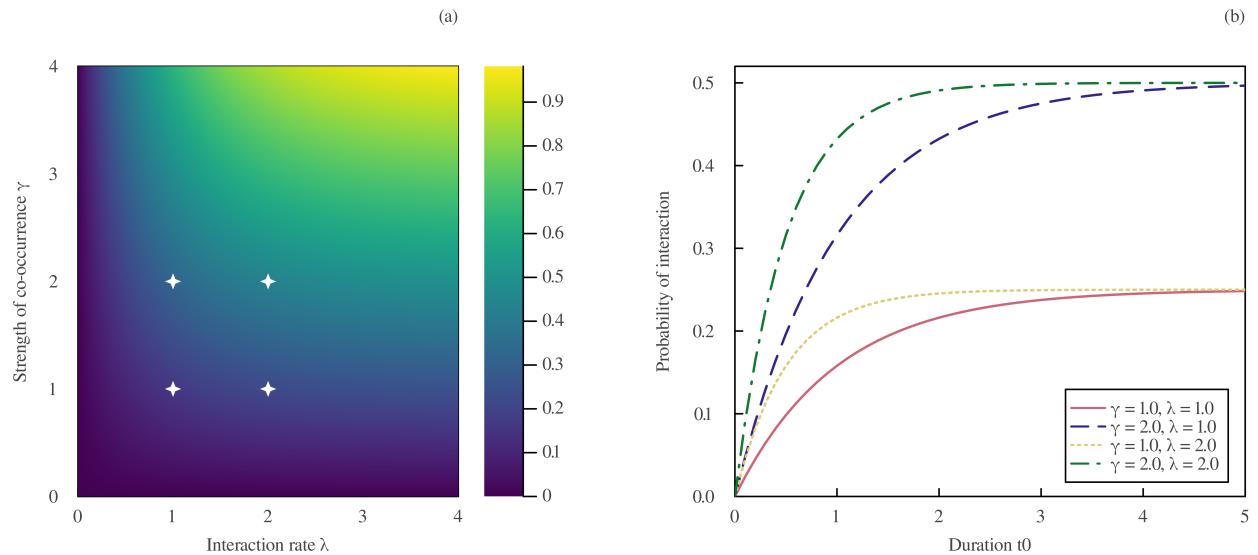


Figure 3: **Parameters of the spatiotemporally explicit model of interactions.** (a) Probability of local interaction given by the process model (eq. 11) under different values of λ and γ , with $t_0 = 1$. The parameter values used in the right panel are denoted by the white stars. (b) Scaling of the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

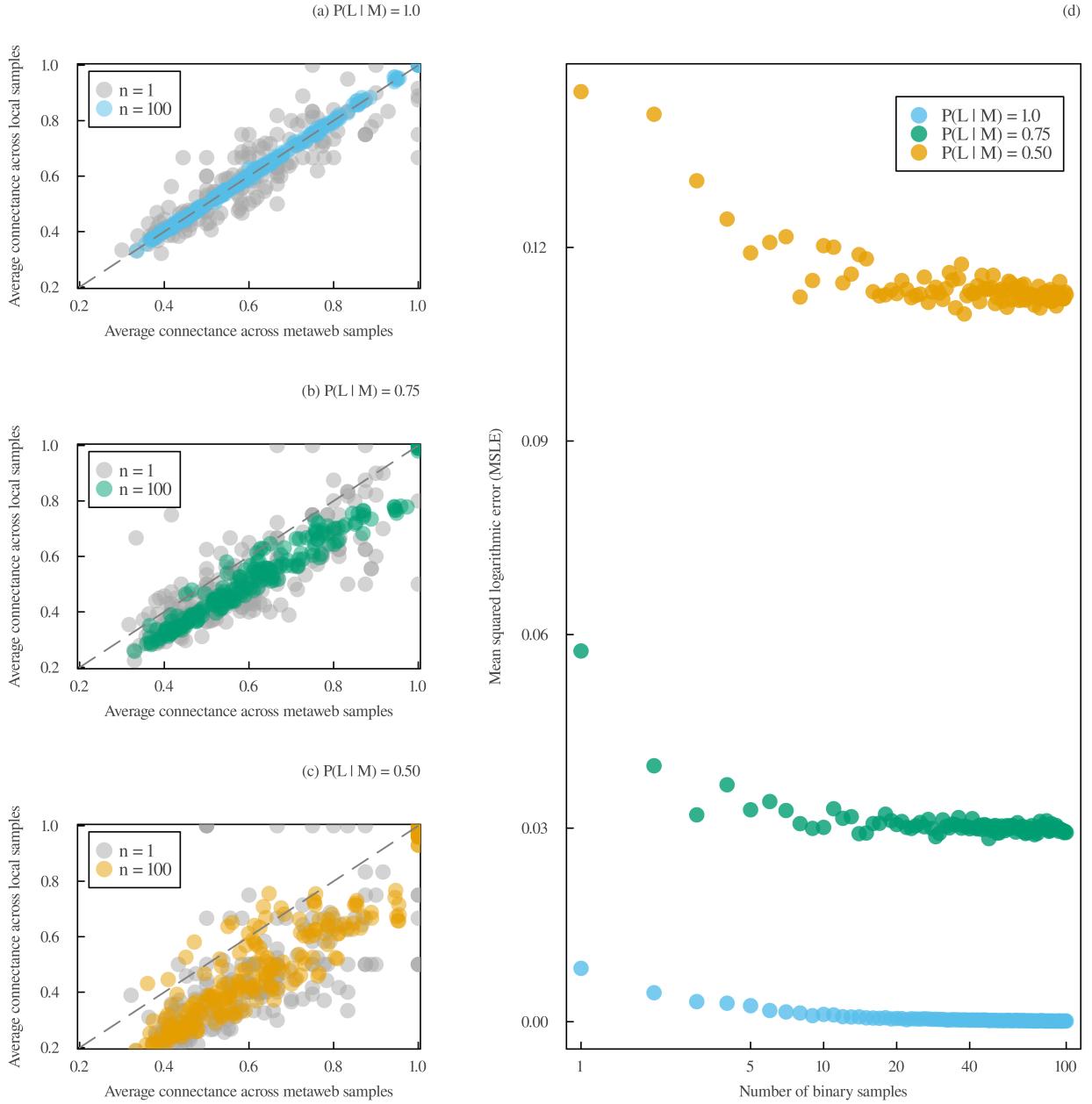


Figure 4: Connectance of sampled binary interaction webs. Comparison between the average connectance of binary interaction network samples obtained from the local and metawebs of probabilistic interactions. Each dot corresponds to a different site. The local probability of interaction between potentially 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.