

The biological interpretation of probabilistic food webs

Francis Banville^{1,2,3,‡} Tanya Strydom^{1,3,‡} Timothée Poisot^{1,3}

¹ Université de Montréal ² Université de Sherbrooke ³ Quebec Centre for Biodiversity Science

‡ These authors contributed equally to the work

Correspondance to:

Francis Banville — francis.banville@umontreal.ca

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



Last revision: *February 12, 2023*

Community ecologists are increasingly thinking probabilistically when it comes to food webs and other ecological networks. Assuredly, the benefits of representing ecological interactions as probabilistic events (e.g., how likely are species to interact?) instead of binary objects (e.g., do species interact?) are numerous, from a better assessment of the spatiotemporal variation of trophic interactions to an increase capacity to reconstruct networks from sparse data. However, probabilities need to be used with caution when working with species interactions. Indeed, depending on the system at hand and the method used to build probabilistic networks, probabilities can have different interpretations that imply different ways to manipulate them. At the core of these differences lie the distinction between assessing the likelihood that two groups of individuals *can* interact and the likelihood that they *will* interact. This impacts the spatial, temporal, and taxonomic scaling of interaction probabilities, thus enlightening the need to properly define them in their ecological context. Published data on probabilistic species interactions are poorly documented in the literature, which impedes our ability to use them appropriately. With these challenges in mind, we propose a general approach to thinking about probabilities in regards to ecological interactions, with a strong focus on food webs, and call for better definitions and conceptualizations of probabilistic ecological networks, both at the local and regional scales.

1 Introduction

2 Cataloging species interactions across space is a gargantuan task. At the core of this challenge lies the
3 spatiotemporal variability of ecological networks (Poisot *et al.* 2012, 2015), which makes documenting the
4 location and timing of interactions difficult. Indeed, it is not sufficient to know that two species have the
5 biological capacity to interact to infer the realization of their interaction at a specific time and space
6 (Dunne 2006). Taking food webs as an example, a predator species and its potential prey must first
7 co-occur in order for a trophic interaction to take place (Blanchet *et al.* 2020). They must then encounter,
8 which is conditional on their relative abundances in the ecosystem and the matching of their phenology
9 (Poisot *et al.* 2015). Finally, the interaction occurs only if the predators have a desire to consume their prey
10 and are able to capture and ingest them (Pulliam 1974). Environmental (e.g. temperature and presence of
11 shelters) and biological (e.g. physiological state of both species and availability of other prey species)
12 factors contribute to this variability by impacting species co-occurrence (Araujo *et al.* 2011) and the
13 realization of their interactions (Poisot *et al.* 2015). In this context, the development of computational
14 methods in ecology can help alleviate the colossal sampling efforts required to document species
15 interactions across time and space (Strydom *et al.* 2021). Having a better portrait of species interactions
16 and the emerging structure of their food webs is important since it lays the groundwork for understanding
17 the functioning, dynamics, and resilience of ecosystems worldwide (e.g., Proulx *et al.* 2005; Pascual *et al.*
18 2006; Delmas *et al.* 2019).

19 The recognition of the intrinsic variability of species interactions and the emergence of numerical
20 methods have led ecologists to rethink their representation of ecological networks, slowly moving from a
21 binary to a probabilistic view of species interactions (Poisot *et al.* 2016). This has several benefits. For
22 example, probabilities represent the limit of our knowledge about species interactions and can inform us
23 about the expected number of interactions and emerging network properties despite this limited
24 knowledge (Poisot *et al.* 2016). They are also very helpful in predicting the spatial distribution of species
25 within networks (Cazelles *et al.* 2016) and the temporal variability of interactions (Poisot *et al.* 2015),
26 generating new ecological data (e.g., Strydom *et al.* 2022), and identifying priority sampling locations of
27 species interactions (see Andrade-Pacheco *et al.* 2020 for an ecological example of a sampling optimization
28 problem). Moreover, the high rate of false negatives in ecological network data, resulting from the
29 difficulty of witnessing interactions between rare species, makes it hard to interpret non-observations of

30 species interactions ecologically (Catchen *et al.* 2023). Using probabilities instead of yes-no interactions
31 accounts for these observation errors; in that case, only forbidden interactions (Jordano *et al.* 2003; Olesen
32 *et al.* 2010) would have a probability value of zero (but see Gonzalez-Varo & Traveset 2016). Many
33 measures have been developed to describe the structure (Poisot *et al.* 2016) and diversity (Ohlmann *et al.*
34 2019; Godsoe *et al.* 2022) of probabilistic interactions, which shows the potential of this framework in the
35 study of a variety of ecological phenomena.

36 However, representing species interactions probabilistically can also be challenging. Beyond
37 methodological difficulties in estimating these numbers, there are important conceptual challenges in
38 defining what we mean by “probability of interactions.” To the best of our knowledge, because the
39 building blocks of this mathematical representation of food webs are still being laid, there is no clear
40 definition found in the literature or data standard when it comes to publishing data on probabilistic
41 interactions (see Salim *et al.* 2022 for a discussion on data standardization for mutualistic networks). This
42 is worrisome, since working with probabilistic species interactions without clear guidelines could be
43 misleading as much for field ecologists as for computational ecologists who use and generate these data.
44 In this contribution, we outline different ways to define and interpret interactions probabilities in network
45 ecology and propose an approach to thinking about them. These definitions mostly depend on the study
46 system (e.g. local network or metaweb) and on the method used to generate them. We show that different
47 definitions can have different ecological implications, especially regarding spatial, temporal, and
48 taxonomic scaling. Although we will focus on food webs, our observations and advice can be applied to all
49 types of ecological networks, from plant-pollinator to host-parasite networks. Indeed, all ecological
50 networks, whether they are unipartite or bipartite, share fundamental commonalities in their biological
51 conceptualization and mathematical representation that support these comparisons (i.e., they all describe
52 groups of individuals interacting with each other). Regardless of the study system, we argue that
53 probabilities should be better documented, defined mathematically, and used with caution when
54 describing species interactions.

55 Definitions and interpretations

56 Food-web representations

57 The basic unit of food webs and other ecological networks are individuals that interact with each others
58 (e.g., by predation; Elton 2001), forming individual-based networks. The aggregation of these individuals
59 into more or less homogeneous groups (e.g., populations, species, trophic species, families) allows us to
60 represent networks at broader scales, which impacts the properties and behaviour of these systems
61 (Guimarães 2020). A network's nodes can thus designate distinct levels of organization, whereas the edges
62 linking these nodes can describe a variety of interaction measures. When using a Boolean (yes-no)
63 representation of biotic interactions, the observation that one individual from group (or node) i interacts
64 with another individual from group j is enough to set the interaction $A_{i,j}$ to 1. This simplified
65 representation of food webs is a highly valuable source of ecological information (Pascual *et al.* 2006) even
66 though it overlooks important factors regarding interaction strengths. These, in turn, can be represented
67 using weighted interactions, which better describe the energy flows, demographic impacts or frequencies
68 of interactions between nodes (Berlow *et al.* 2004; Borrett & Scharler 2019), with $A_{i,j} \in \mathbb{N}$ or \mathbb{R} depending
69 on the measure. For example, they can be used to estimate the average number of prey individuals
70 consumed by the predators in a given time period (e.g., the average number of fish in the stomach of a
71 piscivorous species). Interaction strengths can also be used as good estimators of the parameters
72 describing species interactions in a Lotka-Volterra model (e.g., Emmerson & Raffaelli 2004). This extra
73 amount of ecological information typically comes at a cost of greater sampling effort or data requirement
74 in predictive models (Strydom *et al.* 2021), which can lead to high uncertainties when building these types
75 of networks. Therefore, important methodological and conceptual decisions must be made when
76 sampling and building food webs.

77 The uncertainty and spatiotemporal variability of both types of trophic interactions (Boolean and
78 weighted) can be represented probabilistically. On one hand, Boolean interactions follow a Bernoulli
79 distribution $A_{i,j} \sim \text{Bernoulli}(p)$, with p being the probability of interactions. The only two possible
80 outcomes are the presence ($A_{i,j} = 1$) or absence ($A_{i,j} = 0$) of an interaction between the two nodes.
81 Weighted interactions, on the other hand, can follow various probability distributions depending on the
82 measure used. In this case, the event's outcome is the value of interaction strength. For instance, weights
83 can follow a Poisson distribution $A_{i,j} \sim \text{Poisson}(\lambda)$ when predicting frequencies of interactions between

84 pairs of nodes, with λ being the expected rate of interaction. Note that weighted interactions can be
85 converted to probabilistic interactions by normalizing. The definition and interpretation of parameters
86 like p and λ are inextricably linked to environmental and biological factors such as species relative
87 abundance, traits, area, and time, depending on the type of interaction.
88 Because Boolean species interactions are much more documented in the literature, our primary focus in
89 this contribution will be on addressing the challenges in defining and interpreting p for pairwise species
90 interactions.

91 The first aspect to take into consideration when estimating or interpreting probabilities of interactions is
92 knowing if they describe the likelihood of potential or realized interactions. A potential interaction is
93 defined as the biological capacity of two species to interact (i.e., the probability that they *can* interact)
94 whereas a realized interaction refers to the materialization or observation of this interaction in a delineated
95 space and time period (i.e., the probability that they interact). Here, we will use the terms *metaweb* to
96 designate networks of potential interactions and *local networks* for those of realized interactions. Frequent
97 confusion arises among ecologists over the use of these two terms, especially in a probabilistic context.
98 Indeed, in many studies of probabilistic ecological networks, it remains unclear when authors describe
99 potential or realized interactions, or when so-called probabilities are actually *interaction scores*. Likewise,
100 probabilistic potential interactions are often used as realized interactions (and conversely), even when the
101 type of interaction is clearly indicated. We believe that a better understanding of these differences and
102 concepts would alleviate interpretation errors and help ecologists use these numbers more appropriately.

103 **Probabilistic metawebs**

104 Metawebs are networks of potential interactions, representing the probability that two taxa can interact
105 regardless of biological plasticity, environmental variability or co-occurrence. Instead of describing
106 stochastic biological processes occurring in nature, probabilistic potential interactions can be thought of as
107 a measure of imperfect knowledge about the capacity of two taxa to interact. They are the network
108 analogue to the species pool, where local networks originate from a subset of both species (nodes) and
109 interactions (edges). For this reason, they have been initially designed for broad spatial, temporal, and
110 taxonomic scales (e.g, species food webs at the continental scale). However, in the next section, we argue
111 that this concept can also be used at smaller scales, with similar ecological meaning.

112 We can express the probability that two taxa i and j can interact in a metaweb M as

$$P_M(i \rightarrow j), \quad (1)$$

113 which is context independent. In other words, the probability that two species can interact is not
114 contingent on location, time, and environmental factors. Nevertheless, one aspect of a metaweb that could
115 be conditional on these factors is the list of species (or taxa) it is built from when assembled for a specific
116 region.

117 Starting from a selected set of species, metawebs can be built using different data sources, including
118 literature review, field work, and predictive models (e.g., the metaweb of Canadian mammals inferred by
119 Strydom *et al.* 2022). Every pair of species that has been observed to interact at least once can be given a
120 probability of interaction of 1; we know that they *can* interact. This means that rare interactions can
121 technically be given high probabilities in the metaweb. Unobserved interactions, on the other hand, are
122 given lower probabilities, going as low as 0 for forbidden links. Two important nuances must however be
123 made here. Because of observation errors due to taxonomic misidentifications and ecological
124 misinterpretations (e.g., due to cryptic species and interactions; Pringle & Hutchinson 2020), many
125 observations of interactions are actually false positives. Similarly, forbidden interactions can be false
126 negatives if e.g. they have been assessed for specific phenotypes, locations or time. Implementing a
127 Bayesian framework, which updates prior probabilities of interactions with empirical data, could lessen
128 these errors.

129 **Probabilistic local networks**

130 As opposed to metawebs, probabilistic local food webs represent the likelihood that two species will
131 interact at a specific location and within a given time period; in other words, they are context dependant.
132 They could also represent the likelihood of observing these interactions within a given area and time. To
133 be specific, space is defined here as the geographic coordinates (x, y) of the spatial boundaries delineating
134 the system, whereas time is the time interval t during which interactions were sampled or for which they
135 were predicted. We want to point out that they are not single values, but rather continued dimensions that
136 could be outlined differently depending on the study system. Regardless of how they were defined, they
137 always delineate a specific area A and duration t . These could refer to the sampled area and duration or to

138 the targeted location and time period.

139 Many factors could be taken into consideration when estimating the probability that a predator species i
140 interacts with a given prey species j locally. One of the most important is species co-occurrence C , which
141 is a Boolean describing if both species can be found at location and time (x, y, t) . Surely, the probability
142 that the interaction is realized must be 0 when species do not co-occur ($C = 0$). Interaction probabilities
143 can also be conditional on other biological and environmental variables, such as temperature,
144 precipitation, presence of shelters, phenotypic plasticity, phenology, and presence of other interacting
145 species in the network. These conditions can affect species traits, which greatly impact the capacity of
146 species to interact (Poisot *et al.* 2015). Similarly, species relative abundance is another important predictor
147 of the probability of interaction, because it impacts the probability that species will randomly encounter
148 (Canard *et al.* 2012; Canard *et al.* 2014; Poisot *et al.* 2015). Here, we will use the variable Ω as a substitute
149 for the biological and ecological context in which interaction probabilities were estimated, including the
150 presence of higher-order interactions. This gives us the following equation for the probability of realized
151 interaction between species (or taxa) i and j in a local network N :

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

152 which can be read as the probability of local interaction between the two species in an area A and time
153 interval t , given their co-occurrence C and specific environmental and biological conditions Ω . These
154 conditions do not systematically need to be specified for all studies. However, when they are, they should
155 be made explicit in the metadata.

156 Multiple difficulties of interpretation arise when the conditions are not clearly specified, which we found
157 is often the case in the literature. For example, if $P_N(i \rightarrow j | C = 1)$ represents the probability that two
158 co-occurring species interact (i.e., the edge's probability value), $P_N(i \rightarrow j)$ denotes instead the probability
159 of interaction without knowing if they co-occur (i.e., the product of the nodes and edge's probability
160 values). For practical reasons, probabilistic ecological networks are generally represented as matrices of
161 probabilities (i.e., matrices of edges without node values), whose elements are thus hard to interpret
162 without clear indications about C . Overall, when probabilities of interactions are estimated using specific
163 values of A , t , C , and Ω , ecologists should make them explicit in their metadata, preferably using
164 mathematical equations to avoid any ambiguity. Below, we will see examples of why this matters when it

comes to spatial, temporal, and taxonomic scaling of biotic interactions.

[Table 1 about here]. Articles using probabilistic interactions and the definitions and variables they considered.

Applications of probabilistic interactions data

Inferring probabilistic local food webs from metawebs

Even though the spatiotemporal variability of interactions is not considered in metawebs, they can still be useful to reconstruct local networks of realized interactions. Indeed, local networks are formed from subsets of their metaweb (called subnetworks), which are obtained by selecting a subset of both species and interactions (Dunne 2006). Because a community's composition is arguably easier to sample (or predict) than its interactions, the biggest challenge is to sample links from the metaweb. This becomes a conceptual issue when we consider how potential and realized interactions differ. Despite these concerns, metawebs remain an important source of ecological information that can be leveraged for realistically predicting spatially explicit networks. First, metawebs set the upper limit of species interactions (McLeod *et al.* 2021), i.e. the probability that two species interact at a specific location is always lower or equal to the probability of their potential interaction:

$$P_N(i \rightarrow j|A, t, C, \Omega) \leq P_M(i \rightarrow j). \quad (3)$$

Therefore, inferring local networks from their metaweb keeping the same values of interaction probability would generate systematic biases in the prediction. In that case, these networks would instead be called *spatially explicit* or *local* metawebs (i.e., smaller-scale networks of potential interactions). Second, the structure of local networks is constrained by the one of their metaweb (Saravia *et al.* 2022). This suggests that a metaweb not only constrains the pairwise interactions of its corresponding local networks, but also their emerging properties. Inferring the structure of local networks from the metaweb could thus help estimate more realistically the likelihood that potential interactions are realized and observed locally (Strydom *et al.* 2021).

[Figure 1 about here]. Empirical example of the association between the number of interactions in

189 realized local food webs and the number of interactions in the corresponding subnetworks of their
190 regional metaweb. We should expect the association to be linear below the 1:1 line, illustrating eq. 3.

191 **Sampling random draws from probabilistic food webs**

192 Another conceptual challenge encountered when using probabilistic food webs is the prediction of
193 Boolean networks across space. Lets take $n \times n$ grid cells each representing a probabilistic food web. If
194 they contain potential interactions, a single random trial must be conducted for each pairwise interaction
195 across the region (i.e., we should have only one random realization of the regional metaweb). On the
196 contrary, if they represent probabilities of realized interactions, each food web must be independently
197 sampled (i.e., n^2 independent random draws). This has direct implications on the spatial scaling of
198 interactions. For example, let N_2 be another network of area $A_2 < A_0$ within A_0 and disjoint from N_1 ,
199 such as N_1 and N_2 form N_0 (think of two contiguous cells that together delineate N_0). All other things
200 being equal, we should expect the probability that i and j interacts in A_0 to be
201 $P_{N_0}(i \rightarrow j) = 1 - (1 - P_{N_1}(i \rightarrow j)) \times (1 - P_{N_2}(i \rightarrow j))$ if N_1 and N_2 are independently sampled. This also
202 implies that we should expect interactions to be realized in a certain number of local networks depending
203 on the probability value, which is not the case with metawebs. Note that spatial auto-correlation and the
204 concept of meta-network (i.e., networks of networks) could invalidate the statistical assumption of
205 independence. Nevertheless, the fundamental difference in sampling metawebs and local networks stands
206 even when considering these factors. This difference in sampling further sheds light on the importance of
207 clearly defining interaction probabilities. What we consider as a *Bernoulli trial*, when randomly drawing
208 deterministic networks from probabilistic food webs, depends on our biological interpretation of these
209 probabilities.

210 **Describing the spatial and temporal scaling of probabilistic interactions**

211 Metawebs and local networks intrinsically differ in their relation to scale. On one hand, as mentioned
212 above, probabilistic metawebs are context independent, i.e., probabilistic pairwise interactions do not scale
213 with space and time because they depend solely on the biological capacity of the two taxa to interact. This
214 implies that the estimated likelihood that two species can potentially interact should be the same among
215 all metawebs in which they are present. In practice, this is rarely the case because ecologists use different

methods and data to estimate these probabilities of interactions (e.g., different sampling area and time period). However, in the case where local metawebs $M_{x,y}$ are subsampled from their regional counterpart M_0 , we should expect edge values to be identical among all networks, regardless of their spatial scale, i.e. $P_{M_{x,y}}(i \rightarrow j) = P_{M_0}(i \rightarrow j)$. On the other hand, local probabilistic networks are indissociable from their spatial and temporal contexts because there are more opportunities of interactions (e.g., more individuals, more trait variations, more chance of encounter) in a larger area and longer time period (McLeod *et al.* 2020). Let N_0 be a local probabilistic food web delineated in an area A_0 and N_1 a network of area $A_1 < A_0$ within A_0 . We should expect the probability that i and j interacts to scale spatially, i.e. $P_{N_1}(i \rightarrow j | A < A_0) \leq P_{N_0}(i \rightarrow j | A = A_0)$. Similarly, the temporal scaling of probabilistic local food webs could be manifested through the effect of sampling effort on the observation of interactions (Jordano 2016; McLeod *et al.* 2021) or of time itself on their realization (Poisot *et al.* 2012). There are many network-area relationships (e.g., Wood *et al.* 2015; Galiana *et al.* 2018) and interaction accumulation curves (e.g. Jordano 2016) explored in the literature. These could inspire the development and testing of different equations describing the spatiotemporal scaling of probabilistic pairwise interactions in local networks, which are over the scope of this manuscript.

[Figure 2 about here]. Conceptual figure showing (1) the spatiotemporal scaling of probabilistic metawebs and local food webs and (2) the spatial sampling of metawebs and local food webs into Boolean networks.

Making probabilistic interactions spatiotemporally explicit

The variability of species interactions spurred the development of methods aiming at predicting ecological networks at fine spatial and temporal scales. For example, Bohan *et al.* (2017) proposed a framework to reconstruct networks in real time using continuous biomonitoring. Here, we will build on these studies by proposing a simple model to make probabilistic local networks spatiotemporally explicit. These types of models could prove useful when inferring food webs across time and space from sparse data. However, they are not suitable for metawebs, which are static objects.

One way that probabilistic food webs can be made spatiotemporally explicit is by modelling interactions between co-occurring species as a Poisson process with rate λ . Specifically, if the total observation time for a location is t_0 , the probability that two co-occurring species i and j will interact during this time period is $P_N(i \rightarrow j | C_{i,j} = 1, t = t_0) = 1 - e^{-\lambda t_0}$, which approaches 1 when $t \rightarrow \infty$. The value of the parameter λ

could be estimated using prior data on interaction strengths, if available. Additionally, we can estimate the probability of co-occurrence at location (x, y) with $P_{x,y}(C_{i,j} = 1) = P_{x,y}(i)P_{x,y}(j)\gamma$, where $P_{x,y}(i)$ and $P_{x,y}(j)$ are respectively the probabilities of occurrence of species i and j and γ is the strength of association between occurrence and co-occurrence, as defined in Cazelles *et al.* (2016). Note that in empirical networks, γ is typically > 1 (Catchen *et al.* 2023). The observation of this interaction would thus follow a Bernoulli distribution with parameter $p = p_A(x, y)p_B(x, y)\gamma(1 - e^{-\lambda t_0})$. This simple model could be customized in many ways, e.g. by linking λ with given environmental variables or by adding in observation errors (i.e., probability of false negatives and false positives; Catchen *et al.* (2023)).

Exploring different levels of organization

How do interaction probabilities scale taxonomically?

- There are different biological interpretations of probabilities in food webs at the individual level and at higher taxonomic levels.
- How does the scaling up of the nodes from an individual to population to any higher taxonomic group change our interpretation of interaction probabilities? How does the aggregation change our interpretation?
- Why would we want to scale networks taxonomically?
- Do all nodes need to be the same taxonomic scale, within a network?
- How is it similar and different to spatial and temporal scaling? Basically, all kinds of scaling are just different ways to aggregate individuals or nodes.
- Papers: Vázquez *et al.* (2022)

[Figure 3 about here]. Conceptual figure of how a scale up of the nodes from an individual to a population to any higher taxonomic group change our interpretation of the probability of interaction.

Conclusion

The emergence of probabilistic thinking in network ecology has paved the way to a better assessment of the spatiotemporal variability and uncertainty of biotic interactions. However, measuring probabilities

empirically can be strenuous given the difficulties of deciphering species and interactions (Pringle & Hutchinson 2020). In this context, the development of computational methods makes it possible to estimate interaction probabilities at large scales, which in turn can pinpoint where we should go to optimise our sampling effort for better resolving local food webs.

In this contribution, we showed that network metadata are perhaps as important as interaction data themselves when it comes to interpreting probabilistic food webs in ecological terms. First, the type of probabilistic interaction (potential or realized) impacts the importance of scale, with interactions in metawebs being scale independent, both spatially and temporally. Second, the conditions in which local networks were estimated (e.g., area, time, biological and environmental factors) and the attributes of the interacting species that were considered (e.g., species co-occurrence) are essential contextual factors that impact the mathematical representation of probabilities and their resulting behaviour. Third, the biological interpretation of probabilities changes with the level of organization of the network due to the aggregation of individuals into broader groups. All these information should be available as clear metadata so that ecologists can use probabilistic network data appropriately.

Moreover, many statistical models in ecology that yield accurate predictions of biotic interactions are black boxes difficult to interpret. Ecologists should be careful before using the output of these models as probabilistic objects, since there is often a thin line between a real probability and a non-probabilistic predictive number (or score). Probabilities are numbers between 0 and 1 that sum to 1 and either represent the expected frequency of a phenomenon or the degree of belief that it will be realized. Non-probabilistic scores, which are more akin to interaction weights, have different mathematical properties, which impacts how we should handle these numbers in a spatially or temporally explicit context. Therefore, researchers should use their expertise to assess if their interaction data are actually probabilities or scores. This should also be added to the metadata before sharing them, as well as the methods used to build the networks.

Better metadata documentation would allow researchers to use and manipulate probabilistic ecological interactions according to how they were actually defined and obtained. This would support better scientific practices, in particular when these data are used for ecological prediction and forecasting. For instance, we showed that building a rigorous workflow to predict local networks from a probabilistic metaweb requires a good understanding of the data at hand. Similarly, explicitly stating the context in which probabilistic data were estimated would help using forecasting food-web models more rigorously under specific climate change and habitat use scenarios. Regardless of the method and application,

299 fostering a better foundation for probabilistic reasonings in network ecology, from the very nature of
300 probabilities and biotic interactions, is essential.

301 **Acknowledgement**

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

309 **References**

- 310 Andrade-Pacheco, R., Rerolle, F., Lemoine, J., Hernandez, L., Meïté, A., Juziwelo, L., *et al.* (2020). Finding
311 hotspots: Development of an adaptive spatial sampling approach. *Scientific Reports*, 10, 10939.
- 312 Araujo, M.B., Rozenfeld, A., Rahbek, C. & Marquet, P.A. (2011). Using species co-occurrence networks to
313 assess the impacts of climate change. *Ecography*, 34, 897–908.
- 314 Berlow, E.L., Neutel, A.-M., Cohen, J.E., De Ruiter, P.C., Ebenman, B., Emmerson, M., *et al.* (2004).
315 Interaction strengths in food webs: Issues and opportunities. *Journal of Animal Ecology*, 73, 585–598.
- 316 Blanchet, F.G., Cazelles, K. & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions.
317 *Ecology Letters*, 23, 1050–1063.
- 318 Bohan, D.A., Vacher, C., Tamaddon-Nezhad, A., Raybould, A., Dumbrell, A.J. & Woodward, G. (2017).
319 Next-Generation Global Biomonitoring: Large-scale, Automated Reconstruction of Ecological
320 Networks. *Trends in Ecology & Evolution*, 32, 477–487.
- 321 Borrett, S.R. & Scharler, U.M. (2019). Walk partitions of flow in Ecological Network Analysis: Review and
322 synthesis of methods and indicators. *Ecological Indicators*, 106, 105451.

Canard, E.F., Mouquet, N., Mouillot, D., Stanko, M., Miklisova, D. & Gravel, D. (2014). Empirical Evaluation of Neutral Interactions in Host-Parasite Networks. *The American Naturalist*, 183, 468–479.

Canard, E., Mouquet, N., Marescot, L., Gaston, K.J., Gravel, D. & Mouillot, D. (2012). Emergence of Structural Patterns in Neutral Trophic Networks. *PLOS ONE*, 7, e38295.

Catchen, M.D., Poisot, T., Pollock, L.J. & Gonzalez, A. (2023). The missing link: Discerning true from false negatives when sampling species interaction networks.

Cazelles, K., Araujo, M.B., Mouquet, N. & Gravel, D. (2016). A theory for species co-occurrence in interaction networks. *Theoretical Ecology*, 9, 39–48.

Delmas, E., Besson, M., Brice, M.-H., Burkle, L.A., Dalla Riva, G.V., Fortin, M.-J., *et al.* (2019). Analysing ecological networks of species interactions. *Biological Reviews*, 94, 16–36.

Dunne, J.A. (2006). The Network Structure of Food Webs. In: *Ecological networks: Linking structure and dynamics* (eds. Dunne, J.A. & Pascual, M.). Oxford University Press, pp. 27–86.

Elton, C.S. (2001). *Animal Ecology*. University of Chicago Press, Chicago, IL.

Emmerson, M.C. & Raffaelli, D. (2004). Predator-prey body size, interaction strength and the stability of a real food web. *Journal of Animal Ecology*, 73, 399–409.

Galiana, N., Lurgi, M., Claramunt-López, B., Fortin, M.-J., Leroux, S., Cazelles, K., *et al.* (2018). The spatial scaling of species interaction networks. *Nature Ecology & Evolution*, 2, 782–790.

Godsoe, W., Murray, R. & Iritani, R. (2022). Species interactions and diversity: A unified framework using Hill numbers. *Oikos*, n/a, e09282.

Gonzalez-Varo, J.P. & Traveset, A. (2016). The Labile Limits of Forbidden Interactions. *Trends in Ecology & Evolution*, 31, 700–710.

Guimarães, P.R. (2020). The Structure of Ecological Networks Across Levels of Organization. *Annual Review of Ecology, Evolution, and Systematics*, 51, 433–460.

Jordano, P. (2016). Sampling networks of ecological interactions. *Functional Ecology*, 30, 1883–1893.

Jordano, P., Bascompte, J. & Olesen, J.M. (2003). Invariant properties in coevolutionary networks of plant-animal interactions. *Ecology Letters*, 6, 69–81.

349 McLeod, A., Leroux, S.J., Gravel, D., Chu, C., Cirtwill, A.R., Fortin, M.-J., *et al.* (2021). Sampling and
 350 asymptotic network properties of spatial multi-trophic networks. *Oikos*, n/a.

351 McLeod, A.M., Leroux, S.J. & Chu, C. (2020). Effects of species traits, motif profiles, and environment on
 352 spatial variation in multi-trophic antagonistic networks. *Ecosphere*, 11, e03018.

353 Ohlmann, M., Miele, V., Dray, S., Chalmandrier, L., O'Connor, L. & Thuiller, W. (2019). Diversity indices
 354 for ecological networks: A unifying framework using Hill numbers. *Ecology Letters*, 22, 737–747.

355 Olesen, J.M., Bascompte, J., Dupont, Y.L., Elberling, H., Rasmussen, C. & Jordano, P. (2010). Missing and
 356 forbidden links in mutualistic networks. *Proceedings of the Royal Society B: Biological Sciences*, 278,
 357 725–732.

358 Pascual, M., Dunne, J.A. & Dunne, J.A. (2006). *Ecological Networks: Linking Structure to Dynamics in Food*
 359 *Webs*. Oxford University Press, USA.

360 Poisot, T., Canard, E., Mouillot, D., Mouquet, N. & Gravel, D. (2012). The dissimilarity of species
 361 interaction networks. *Ecology Letters*, 15, 1353–1361.

362 Poisot, T., Cirtwill, A.R., Cazelles, K., Gravel, D., Fortin, M.-J. & Stouffer, D.B. (2016). The structure of
 363 probabilistic networks. *Methods in Ecology and Evolution*, 7, 303–312.

364 Poisot, T., Stouffer, D.B. & Gravel, D. (2015). Beyond species: Why ecological interaction networks vary
 365 through space and time. *Oikos*, 124, 243–251.

366 Pringle, R.M. & Hutchinson, M.C. (2020). Resolving Food-Web Structure. *Annual Review of Ecology,*
 367 *Evolution, and Systematics*, 51, 55–80.

368 Proulx, S.R., Promislow, D.E.L. & Phillips, P.C. (2005). Network thinking in ecology and evolution. *Trends*
 369 *in Ecology & Evolution*, SPECIAL issue: BUMPER book REVIEW, 20, 345–353.

370 Pulliam, H.R. (1974). On the Theory of Optimal Diets. *The American Naturalist*, 108, 59–74.

371 Salim, J.A., Saraiva, A.M., Zermoglio, P.F., Agostini, K., Wolowski, M., Drucker, D.P., *et al.* (2022). Data
 372 standardization of plantpollinator interactions. *GigaScience*, 11, giac043.

373 Saravia, L.A., Marina, T.I., Kristensen, N.P., De Troch, M. & Momo, F.R. (2022). Ecological network
 374 assembly: How the regional metaweb influences local food webs. *Journal of Animal Ecology*, 91,
 375 630–642.

376 Strydom, T., Bouskila, S., Banville, F., Barros, C., Caron, D., Farrell, M.J., *et al.* (2022). Food web
377 reconstruction through phylogenetic transfer of low-rank network representation. *Methods in Ecology*
378 *and Evolution*, 13.

379 Strydom, T., Catchen, M.D., Banville, F., Caron, D., Dansereau, G., Desjardins-Proulx, P., *et al.* (2021). A
380 roadmap towards predicting species interaction networks (across space and time). *Philosophical*
381 *Transactions of the Royal Society B-Biological Sciences*, 376, 20210063.

382 Vázquez, D.P., Peralta, G., Cagnolo, L., Santos, M. & Igual, E. autores contribuyeron por. (2022). Ecological
383 interaction networks. What we know, what we don't, and why it matters. *Ecología Austral*, 32, 670–697.

384 Wood, S.A., Russell, R., Hanson, D., Williams, R.J. & Dunne, J.A. (2015). Effects of spatial scale of
385 sampling on food web structure. *Ecology and Evolution*, 5, 3769–3782.