A quantitative framework for investigating the reliability of empirical network construction: supplemental information

Alyssa R. Cirtwill[1†], & Anna Eklöf[1], Tomas Roslin[2], Kate Wootton[2], Dominique Gravel[3]

[1]Department of Physics, Chemistry
and Biology (IFM)
Linköping University
Linköping, Sweden

[2] Department of Ecology
P.O. Box 7044, Swedish University of Agricultural Sciences
SE-750 07 Uppsala, Sweden

[3] Département de biologie
Université de Sherbrooke
Sherbrooke, Canada

[†] Corresponding author:
alyssa.cirtwill@gmail.com
+46 723 158464

Appendix S1: Simulating the effects of uncertainty on network structure

To demonstrate the effects of nested interaction, process, and detection uncertainty on network structure (see Fig. 1, main text), we can observe how the structure of randomly-generated networks changes as we add layers of uncertainty. We began by constructing 100 "true" networks with 50 consumer species, 50 resource species, and a connectance of 0.1. We ensured that each network had at least one link per consumer and per resource. For each network, we calculated the mean number of links per consumer, mean number of links per resource, and nestedness (NODF) as described in the main text. For simplicity, we assume that there is no interaction uncertainty in this system (i.e., all interactions are either 100% feasible or 100% unfeasible based on species' traits).

Taking each "true" network as a starting point, we then applied process uncertainty. The probability p_{ij} of the occurrence of each interaction on a given sampling day was assigned a random probability drawn from the uniform distribution. As the probability that an interaction will occur during sampling also depends on the probability of the two species co-occurring, we also assigned a random number of sampling days on which each species co-occurred. The number of co-occurrences, n_{ij}, was drawn from a uniform distribution and could be any integer between 0 and 30 (our simulated number of sampling days), inclusive. The number of occurrences, k_{ij}, of each interaction was then drawn from a binomial distribution based on n_{ij} and p_{ij}. If the number of occurrences during sampling was greater than 0 and the interaction occurred in the true network, we included the interaction in the "process-filtered" network. As the inclusion of interactions in the process matrices is random, we created 100 process-filtered networks for each true network. For each network, we calculated the number of consumers and resources (after removing any species with no interactions), mean links per consumer and per resource, connectance, and NODF.

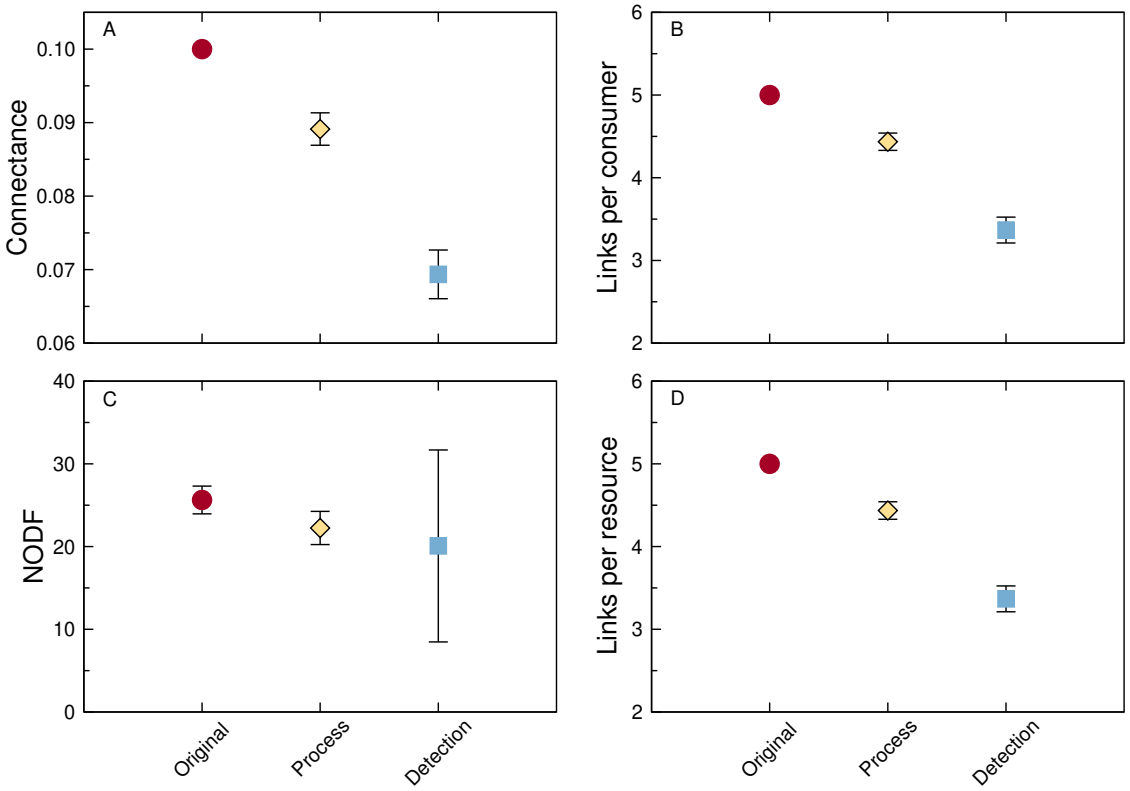
Next, we applied a layer of detection uncertainty to create "detection-filtered" networks. As with the process-filtered networks, we randomly selected the number of times each species co-occurred, n_{ij}, from a uniform distribution. We also randomly selected the probability, p, of detecting each interaction. To reflect the fact that some interactions are easy to observe but many will be more difficult (because one or both interaction partners are rare, cryptic, or difficult to rear in the lab), we drew p from a beta distribution with shape parameters 2 and 10. We created 100 detection-filtered networks for each process-filtered network (10000 per true network). For each network, we calculated the number of consumers and resources (after removing any species with no interactions), mean links per consumer and per resource, connectance, and NODF. Code used in this simulation example is contained in the file *simulation_example.R*, uploaded separately.

With each additional layer of uncertainty, the number of consumers and resources included in the web decreased (mean 49.8 of each species type in detection-filter webs and mean 48.6 of each type in process-filtered webs). The connectance and number of links per consumer and per resource decreased significantly at each level (Fig. S1A,B,D; $p<0.001$ in all cases based on one-tailed Student's t tests). The effects of uncertainty on nestedness were less predictable, particularly in the detection-filtered webs where variability was much

higher than in other web types. Nevertheless, mean nestedness was lower in detection-filtered webs than in the original webs (t=19.9, p<0.001) and lower in the process-filtered webs than the original and detection-filtered webs (Fig. S1C; t=32.8, p<0.001 and t=93.7, p<0.001). We note that these tests may reflect the large number of detection-filtered webs generated by our simulation. A single observed network drawn from these detection-filtered webs might well have a higher nestedness than the "true" network. This simulation demonstrates the fact that process and detection uncertainty mean that we are very unlikely to observe all of the interactions that can truly occur in a system. It also highlights the sometimes unpredictable effects that this uncertainty can have on higher-order structures such as nestedness.

Note that this simulation uses several simplifying assumptions. For instance, we treat interactions as independent and ignore the possibility for, for example, competition between resources. A thorough exploration of the effects of interaction contingency, different distributions of co-occurrences, etc. would be an interesting topic for future work but is beyond the scope of the current study. We also neglect interaction uncertainty in this example. Interaction uncertainty could be included by, for example, modelling interactions between individuals of each species. In this framework, individuals could vary in some trait (e.g., body size) that affects the interactions in which they participate. Very small or very large individuals might not participate in some interactions that are feasible for most members of the species.

Figure S1: We created 100 random networks containing 50 consumers and 50 resources, with a connectance of 0.1. For each of these "original" networks, we created 100 process-filtered ("process") networks incorporating uncertainty around the occurrence of each interaction during sampling. For each "process" network, we further created 100 detection-filtered ("detection") networks incorporating uncertainty around the detection of each interaction. For each network, we show the mean (+/- SD) connectance, links per consumer, nestedness (NODF), and links per resource.



Appendix S1-A: Attributions for PhyloPic images used in Figure 1 (main text)

Table S1: Attributions for silhouette images taken from PhyloPic.org and used in Figure 1 (*main text*). Terms for the CC BY-NC-SA 3.0 license can be found at *https://creativecommons.org/licenses/by-nc-sa/3.0/*. Terms for the CC BY 3.0 license can be found at *https://creativecommons.org/licenses/by/3.0/*.

Image	PhyloPic taxon	Attribution
Lion	Panthera leo	Public domain (uncredited)
Leopard	Panthera pardus	Lukasiniho (CC BY-NC-SA 3.0; image not changed)
Jackal	Canidae	Brian Gratwicke (photo) and T. Michael Keesey (vectorization). (CC BY 3.0; image not changed)
Elephant	Loxodonta africana	Jan A. Venter, Herbert H. T. Prins, David A. Balfour & Rob Slotow (vectorized by T. Michael Keesey) (CC BY 3.0; image not changed)
Zebra	Equus quagga	Cathy (CC BY-NC-SA 3.0; image not changed)
Impala	Aepyceros melampus	Public domain (unknown author)
Hare	Leporidae	Sarah Werning (CC A 3.0; image not changed)
Acacia	Acacia	Mattia Menchetti (public domain)
Grass		Not sourced from PhyloPic. Public domain.

Appendix S2: Can we reduce uncertainty by sampling more?

The absence of an interaction on a given day could be because it was not feasible $[P(L_{ij}) = 0]$, was feasible but did not occur because of local conditions during sampling $[(P(X_{ij}|L_{ij} = 1) = 0]$, or was feasible and occurred during sampling but could not be observed $[P(D_{ij}|X_{ij} = 1, L_{ij} = 1) = 0]$. Note that some external factors can affect more than one level of uncertainty. For example, rare species are less likely to interact at a given site due to neutral processes (Jordano, 2016; Graham & Weinstein, 2018). Interactions involving rare species are also less likely to be detected, unless more sampling effort is focused on rare than common species (Bartomeus, 2013; Jordano, 2016). Thus, abundance affects both the probability of an interaction occurring at a particular site *and* our ability to observe the interaction.

When considering a metaweb, we wish to separate the unobserved interactions where $L_{ij} = 0$ (i.e., unfeasible interactions) from interactions which could occur but did not at a particular sampling site or which occurred but were not detected (Fig. 1, main text). Unfortunately, this is not possible for most interactions (except those where we know an interaction is truly not feasible). That is, we generally do not know *why* an interaction was not observed, only that it was not. An empirical ecologist will instead measure the marginal probability $P(L_{ij}) = k_{ij}/n_{ij}$, where k_{ij} is the number of observed interactions between species i and j and n_{ij} the number of observed co-occurrences. Given this information, the question becomes: how can we reduce the uncertainty around our estimated interaction probability?

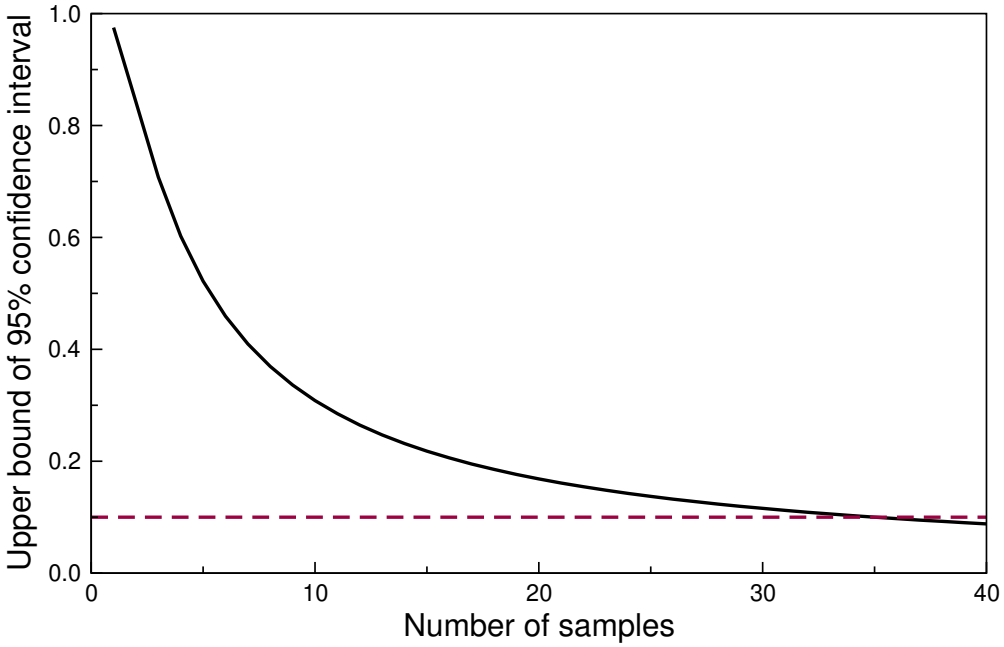
The obvious rule of thumb to reduce uncertainty is "sample more". Where resources permit, increasing sample sizes will reduce uncertainty about the upper bound of the interaction probability and will also increase the probability of detecting unlikely or cryptic interactions (e.g., interactions where $L_{ij} = 1$ but process or detection uncertainty is high). Despite these benefits, we note that there are limits to the utility of increased sampling. Since the probability of observing the co-occurrence of two species will always be higher than the probability of observing their interaction (since the probability of interaction is conditional on both interaction partners being present (Gravel *et al.*, 2018)), we will accumulate observations of co-occurrences faster than we will accumulate observations of interactions (Fig. 2C, main text). While this improves our understanding of the set of species present in the community, it introduces yet more uncertainty into the matrix of feasible interactions. Note that it is better to identify the set of potential interactions than to miss a species, so sampling effort should not be reduced in order to exclude rare species.

An added complication is that increased sampling will not reduce uncertainty evenly across interactions. To record an interaction between species i and j, we need to identify both partners correctly (a non-trivial problem in many food webs; e.g. Kaartinen & Roslin, 2011; Roslin & Majaneva, 2016). For both molecular and rearing techniques, certain types of interactions may go unnoticed due to technical challenges (Wirta *et al.*, 2014). This can bias the set of recorded interactions towards those that are easier to observe (Carstensen *et al.*, 2014; Jordano, 2016). Combining multiple sampling types can reduce this type of bias (Wirta *et al.*, 2014; Jordano, 2016). Nevertheless, despite substantial improvements in

recent decades, most available interaction networks are still undersampled and, moreover, are only subsets of larger communities (Bartomeus, 2013; Jordano, 2016).

In short, high-quality sampling is an essential component of ecological network construction and increasing sampling (length of time, number of sites, number of distinct methods, etc.) can reduce some of the uncertainty about unobserved interactions. We believe, however, that empirical researchers generally sample to the maximum extent possible given constraints of time, resources, and study sites. This practical constraint means that reducing one source of uncertainty is likely to increase another source (e.g., sampling intensively at a single site may reduce detection uncertainty but will increase uncertainty due to variation between sites). The number of observed co-occurrences needed to believe that an unobserved interaction cannot occur is very high (about 35; Fig. S2), and most datasets will not include this level of sampling for all pairs of species (Bartomeus, 2013). For pairs of species with fewer observations, we can reduce uncertainty by combining the observed data available with prior knowledge about the system. This is straightforward to do using a Bayesian approach, as detailed in the main text.

Figure S2: Imperfect detection of interactions increases the number of samples required to be confident that an interaction never occurs. Assume that two species i and j cannot interact (i.e., $\lambda_{ij} = 0$) and that the number of observed interactions follows a binomial distribution depending upon the number of observed interactions k_{ij} and observed co-occurrances n_{ij}. The maximum likelihood estimate (MLE) for λ_{ij} is $\lambda_{ij,MLE} = \frac{k_{ij}}{n_{ij}}$. Note that λ_{ij} is not a point estimate but rather a random variable with an unknown distribution. This means that if $k_{ij} = 0$ in a given sample, this does not necessarily imply that the two species will never interact. Rather, $k_{ij} = 0$ implies that 'no interaction' is the most likely outcome when the species do co-occur but there is nonetheless some chance that the two species *could* interact. Importantly, we can estimate the variance of λ_{ij} as well as the MLE; available methods include the *Wilson score interval*, the *Clopper-Pearson interval*, and the *Agresti-Coull interval* (for details, see [Brown *et al.*, 2001]). The distribution of the MLE for λ_{ij} thus obtained is a key part of the Bayesian distribution for λ_{ij}. Here we show the upper bound (solid black line) of a 95% Clopper-Pearson true credible interval for the interaction probability λ_{ij} for a pair of species that has been observed co-occurring n_{ij} times but never interacting ($k_{ij} = 0$). The upper limit of the credible interval only reaches 0.1 (dashed, red line) with 35 observations. Thus, adding more observations is useful in controlling uncertainty, but the number of observations added must be very high unless we have some other reason to believe that an interaction cannot occur (i.e., a strong prior).



Appendix S3: Mathematical framework

A naive quantification of uncertainty

We start by considering how to naively quantify an interaction probability and its associated uncertainty *for an interaction that has not yet been observed*. Consider the case where a pair of species have been observed co-occurring n_{ij} times, of which they have been observed to interact in $k_{ij} = 0$ cases. To evaluate the uncertainty of this interaction, consider the occurrence of an interaction as a Bernoulli trial. In this framework, the number of successes k_{ij} over n_{ij} trials will follow a binomial distribution:

$$L \sim Bin(n, \lambda), \tag{1}$$

$$P(L = k|\lambda, n) = \binom{n}{k} \lambda^k (1-\lambda)^{n-k}. \tag{2}$$

The parameter λ_{ij}, the probability of observing an interaction over an infinite time interval and area, is the quantity we want to estimate from empirical data. The maximal likelihood estimate (MLE) of λ_{ij} is straightforward to find given k_{ij} and n_{ij}:

$$\lambda_{MLE} = \frac{k}{n}. \tag{3}$$

Posterior distribution of the interaction probability

Here we adopt a Bayesian approach to estimate the posterior distribution of the parameter λ_{ij} :

$$\underbrace{P(\lambda|k,n)}_{Posterior} = \frac{\overbrace{P(k|\lambda,n)}^{Likelihood}\overbrace{P(\lambda)}^{Prior}}{\underbrace{P(k|N)}_{Normaliser}}. \tag{4}$$

According to the above description, the likelihood is simply the binomial distribution (Eq. 2). Since λ_{ij} is a probability, it is bounded between 0 and 1 and the most appropriate prior distribution is the beta:

$$\lambda \sim Beta(\alpha, \beta), \tag{5}$$

which has two shape parameters, α and β.

The beta-binomial distribution is a conjugate distribution of the binomial distribution. This allows us to analytically compute the posterior distribution of a binomial model with a beta prior distribution. We can re-write the posterior distribution of λ_{ij} as:

$$P(\lambda|k,n) = \frac{\lambda^{\alpha+k-1}(1-\lambda)^{\beta+n-k-1}}{B(\alpha+k, \beta+n-k)}, \tag{6}$$

where the function B is the beta function. The posterior distribution of λ_{ij} therefore follows the beta distribution with new parameters $\alpha' = \alpha + k$ and $\beta' = \beta + n - k$.

For interactions which were observed in at least one site (k_{ij}>0), we can set the poseterior probability to 1 as long as the risk of a false positive is low. That is, when constructing a metaweb we can assume that an interaction observed at least once is truly feasible. The case study presented in the main text refers to a plant-galler network compiled by rearing galling insects obtained from galls; in each case, the plant-galler interaction definitely occurred and we are confident in the identification of both species involved (see *Appendix S4* for details). Where the risk of false positives is higher (e.g., plant-pollinator networks where an observed flower visitor may not actually provide a pollination service), it may not be appropriate to assume $\lambda_{ij} = 0$ wherever k_{ij}>0. In such cases, we can model the posterior probabilities of observed interactions just as we do those that were not observed.

Properties of the beta distribution

It is common to perform analyses that require calculating higher-order network properties in interaction networks. The fact that the posterior distribution of λ_{ij} follows a beta distribution makes it straightforward to compute moments and other properties needed for these analyses.

The **average** of λ_{ij} is:

$$\bar{\lambda} = \frac{\alpha + k}{\alpha + \beta + n}, \tag{7}$$

and its **variance** is:

$$Var(\lambda|k) = \frac{(\alpha + k)(\beta + n - k)}{(\alpha + \beta + n)^2(\alpha + \beta + n + 1)} \tag{8}$$

The **mode** of the distribution is:

$$\hat{\lambda} = \frac{\alpha + k - 1}{\alpha + \beta + n - 2}. \tag{9}$$

Appendix S4: Calculating uninformative and informative priors

To create Bayesian distributions for the probability of interactions, we must begin with a prior distribution. This distribution reflects our prior knowledge (or lack thereof) about the likelihood of a given interaction. Priors may be selected to be uninformative, so as not to bias the posterior distribution in any particular direction, or may be informative and based on some prior knowledge about the system. Here we demonstrate how to calculate a prior distribution using a known set of degrees. The degree of a node in a network is defined as its number of connections to other nodes. The degree distribution of a network is then the probability distribution of these degrees over the whole network and the normalised degree (i.e., degree divided by the number of potential interaction partners) could therefore be interpreted as an interaction probability. It is consequently possible to use the degree distribution to inform the prior distribution. The degree distribution could come from several networks, from a similar network (e.g. a known network at slightly different location) or from the network of interest if interaction probabilities for some species are already documented. The latter approach allows researchers to apply information from known, abundant species to the rarest species for which interactions are less frequently documented. That is, we can use detection probabilities for interactions between abundant species to set bounds on detection probabilities for interactions involving rare species.

If our focal network describes a system similar to that in a known network, we can use the distribution of interaction probabilities in that network to inform our prior. The probability of any interaction L_{ij} depends on the degrees of species i and j. Using normalised degrees Δ_i and Δ_j, we can obtain the probability of interaction L_{ij}=$\Delta_i \times \Delta_j$. Similar to the procedure for degree distribution, the distribution of these interaction probabilities can be used to establish a prior distribution before any data from the focal network are collected. For distributions of either degrees or interaction probabilities, the procedure for the estimation of the hyper parameters follows the same approach as described above for connectance except that each measurement is at the individual interaction level instead of the network level.

A quantitative example

The Bayesian framework can be illustrated with a simple quantitative example. Suppose we have $n_{ij} = 10$ observations of co-occurrence between species i and species j in a given time interval and area, and $k_{ij} = 3$ observations of interactions. The maximum likelihood estimate of the interaction probability is simply $\lambda_{ij,MLE} = 3/10 = 0.3$.

Now consider we know that species i is known to interact with 10 species (other than species j), which have the following degrees:

degree=c(14, 4, 2, 3, 17, 6, 2, 15, 1, 1) .

If the network has 20 species total, this gives the normalised degrees:

norm_degree=c(0.65, 0.20, 0.10, 0.15, 0.85, 0.30, 0.10, 0.75, 0.05, 0.05).

Species i has a normalised degree of 0.55 (it interacts with species j and 10 other species). We can combine the normalised degree of i with the normalised degrees of its interaction partners to obtain the following set of interaction probabilities for species i and each of its interaction partners:

int_probs = c(0.358, 0.110, 0.055, 0.082, 0.468, 0.165, 0.055, 0.412, 0.028, 0.028) .

The mean of these interaction probabilities is 0.176, approximately two-thirds the $\lambda_{ij,MLE}$ obtained from the observed data. We can use the distribution of these interaction probabilities as our prior distribution and estimate the uncertainty surrounding our $\lambda_{ij,MLE}$. With some simple R code (function "calculate_parameters", *Appendix S7*), we obtain prior parameters α=0.998 and β=4.63. Using these priors in equations 7 and 8 above (or in the R function "calculate_distribution" in *Appendix S7*), we find a prior $\bar{\lambda}_{ij}$=0.177 and var(λ_{ij})=0.026. Adding the observed data ($n_{ij} = 10$, $k_{ij} = 3$) and using the same code, we obtain posterior parameters α'=4.00 and β'=11.6 and a posterior $\bar{\lambda}_{ij}$=0.256 and var(λ_{ij})=0.012. Comparing the posterior distribution to the prior, we see that the posterior is closer to the observed data and that the additional data about interactions between species i and j has reduced the variance. We may also wish to calculate a credible interval (analogous to the frequentist confidence interval). This is also quite straightforward in R (see function "credible_interval" in *Appendix S7*). In this case, a 95% credible interval for $\bar{\lambda}_{ij}$ is (0.080, 0.491).

Now, consider the case where the two species have never been observed interacting across n_{ij} trials, i-e. $k_{ij} = 0$. The question is then "what is the probability that these two species do not interact"? Since it is not possible to prove that the two species could never interact (strictly speaking, in a Bayesian approach $\lambda_{ij} = 0$ is impossible), we must fix a threshold below which we consider that there is no interaction ($\lambda_{ij} \sim 0$). We call this threshold probability $\lambda_{ij}*$. We then use the cumulative distribution function to estimate $P(\lambda_{ij} < \lambda_{ij} * |L = 0, n)$ for different n_{ij}. The function "samples_for_threshold" in *Appendix S7* calculates distribution function for $\lambda_{ij}*$ with an increasing number of trials. This yields a surprising result: it requires >24 observations of no interactions to be 95% sure that the interaction probability is smaller than $\lambda_{ij}*$=0.1 (recall Fig. S2, Box 2). Note the special case where there is no observation of the two species co-occurring, $n_{ij} = 0$. In this situation, the posterior distribution converges to the prior distribution since the data include no information on the probability with which species might interact should they co-occur.

Appendix S5: Details of *Salix* data collection

The *Salix*-galler-parasitoid meta-network dataset collected by Kopelke *et al.* (2017) consists of a single community type sampled across Europe: willow (*Salix*) species, willow-galling sawflies (Hymenoptera, Tenthredinidae, Nematinae, Euurina), and their natural enemies (hymenopteran parasitoids and coleopteran, lepidopteran, dipteran, and hymenopteran inquilines). The data were collected over 29 years (1982-2010) at 374 unique locations across Europe ranging from Sicily to the Arctic. Since the dataset includes repeated visits to a number of sites, the data set contains 641 site-visits. Here we take the more conservative approach and pool visits to the same site for a sample size of 374 sub-networks.

The meta-network consists of 1,173 different interactions between 52 *Salix* nodes, 92 herbivore nodes, and 126 parasitoid nodes. Interactions were determined by dissecting and rearing gall inhabitants from 165,424 galls. Some sites were visited repeatedly, for a total of 641 site-visits. We consider these to be repeated samples and take the 374 unique sites as our sample size.

Appendix S6: Results for *Salix*-galler component

Computing the prior and posterior distributions

As with the galler-natural enemy component (presented in the main text), we estimated frequencies of *Salix*-galler interactions based on the normalised degree of each species. We obtained prior parameters of α=8.72, β=305 for the *Salix*-galler component. These prior parameters were then used to estimate the posterior distribution.

For species where $n_{ij} = 0$ (3,986/4,992 *Salix*-galler pairs in our dataset), estimates for the mean and variance of λ_{ij} can be calculated directly from the prior distribution. For the *Salix*-galler network, these parameters were: $\bar{\lambda}_{ij}$=0.028, var(λ_{ij})=8.60$\times 10^{-5}$; note that this mean is much lower than that obtained using a prior based on Barbour *et al.* (2016, Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.g7805). Note that, despite the intensive sampling in the **??** dataset, most *Salix*-galler pairs were never observed co-occurring (3,986/4,992; Fig. S3).

For species with $n_{ij} > 0$, the posterior distribution will depend on the values of n_{ij} and k_{ij}. Here we are interested only in cases where $k_{ij} = 0$. This gives α'=α and β'=$\beta + n_{ij}$. For a pair of species which co-occurred at all 374 sites and was never observed to interact, the distribution for the *Salix*-galler network would become $\bar{\lambda}_{ij}$=1.27 $\times$ 10^{-2}, var(λ_{ij})=1.82 $\times$ 10^{-5}. As n_{ij} decreases, the distribution widens and moves away from the origin (Fig. S6). The 95% credible interval for hypothetical *Salix*-galler pairs widened from (0.006, 0.022) if the pair co-occurred at all sites to (0.013, 0.049) if they co-occurred at none (Table S2, *Appendix S11*).

How many samples are required to reach a minimal precision

In our dataset, the entire 95% credible interval for the probability of an interaction between a *Salix*-galler pair was (0.013, 0.049). We may therefore be 95% confident that the interaction probability for *Salix* and galler species that have not been observed co-occurring is below 0.05. As the peak of the prior distribution for the probability of interaction between *Salix* and galler probabilities is around 0.02 (Fig. S6), to be 95% confident that the interaction probability for these species is below 0.01 would require 1029 observed co-occurrences with no interaction - far more than the number of sites in the (Kopelke *et al.*, 2017) dataset.

Scaling up to network metrics

As with the galler-natural enemy component, we calculated posterior probability distributions for *Salix*-galler pairs that were not observed interacting. For each posterior web, we calculated the connectance of each web, as well the mean links per *Salix*, mean links per galler, and nestedness (NODF). Likewise, we created a suite of detection-filtered networks by randomly sampling 99%, 95%, 90%, 80%, 70%, 60%, and 50% of the interactions included in each posterior network. We then calculated the same network properties as described above.

We find, perhaps not surprisingly, that the posterior webs for the Salix-galler network had higher connectances than the original, observed web (C=0.028 for the observed web and $0.082 \leq C \leq 0.096$ for the posterior webs; Fig. S9A). The number of links per *Salix* species in the observed web (L_{Salix}=2.71) was similar to those in the posterior webs ($2.53 \leq L_{Salix} \leq 3.19$; Fig S9B). The number of links per galler, however, was lower in the observed web (L_{galler}=1.47) than in the posterior webs, accounting for the increased connectance ($4.67 \leq L_{galler} \leq 5.88$; Fig. S9D). There was a more substantial difference in the nestedness of the observed and posterior webs: the observed network had NODF=0.560 while the posterior networks were more nested ($1.39 \leq NODF \leq 1.94$). Even the networks sampled with a detection filter of 50% had non-zero nestedness (Fig. S9C). This last result highlights the potential for the possibility for network structure to vary when considering the possibility that unobserved species pairs may interact.

Figure S3: A) Most pairs of *Salix* and gallers were never observed co-occurring despite the high levels of replication in our example dataset. For those pairs that were observed together at least once ($n_{ij} > 0$), the number of observed co-occurrences was generally small (<10). Here we show a histogram of the number of pairs of species observed co-occurring at least once. The 3986 *Salix*-galler pairs never observed co-occurring are omitted from the histogram. **B)** Most pairs of species that were observed at the same site were never observed interacting. Here we show a histogram of the number of observed interactions within pairs of co-occurring species. Species which co-occurred but never interacted are included. **C)** Here we show, for each species pair, the number of observed interactions plotted against the number of observed co-occurrences. *Salix*-galler pairs either are never observed interacting or interact almost every time they co-occur. The red, dashed line indicates a 1:1 relationship between interactions and co-occurrences.

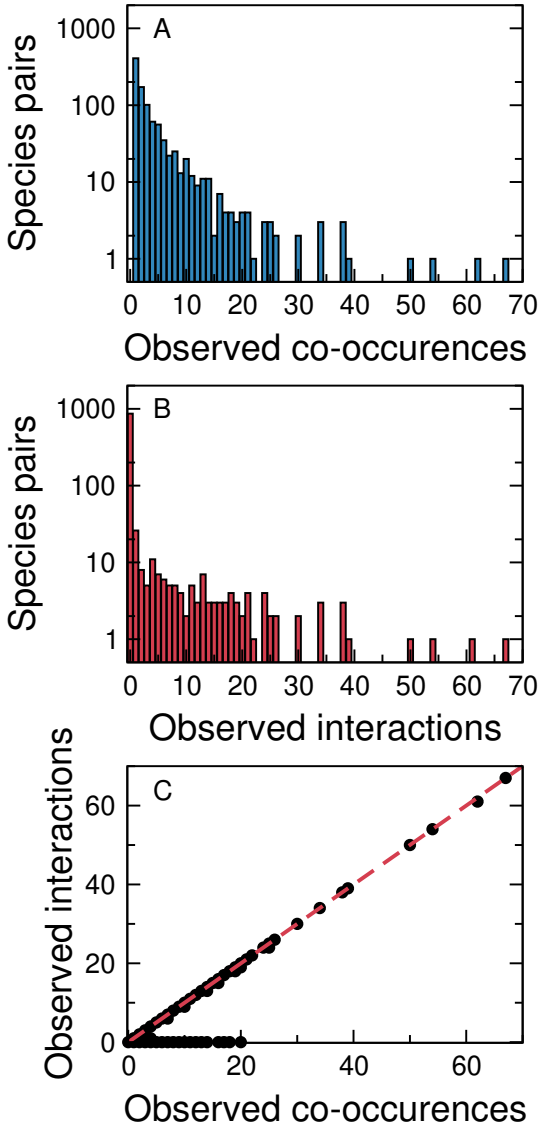


Figure S4: Using prior distributions based on the *Salix*-galler network sampled at a single site in Kopelke *et al.* (2017), we can calculate posterior distributions for the probability of interaction (λ_{ij}) between two species i and j that have not yet been observed interacting. Here we show posterior distributions for λ_{ij} ranging from the prior distribution ($n_{ij} = 0$ observed co-occurrence) to the distribution obtained when the pair of species has been observed co-occurring 100 times. The distribution narrows and approaches zero as the sample size increases. Likewise, the maximum likelihood estimator for the mean probability of interaction (diamonds at top) approaches zero and the 95% credible interval (lines at top) narrows as sample size increases. Dashed lines indicate threshold probabilities of 0.1, 0.05, and 0.01.

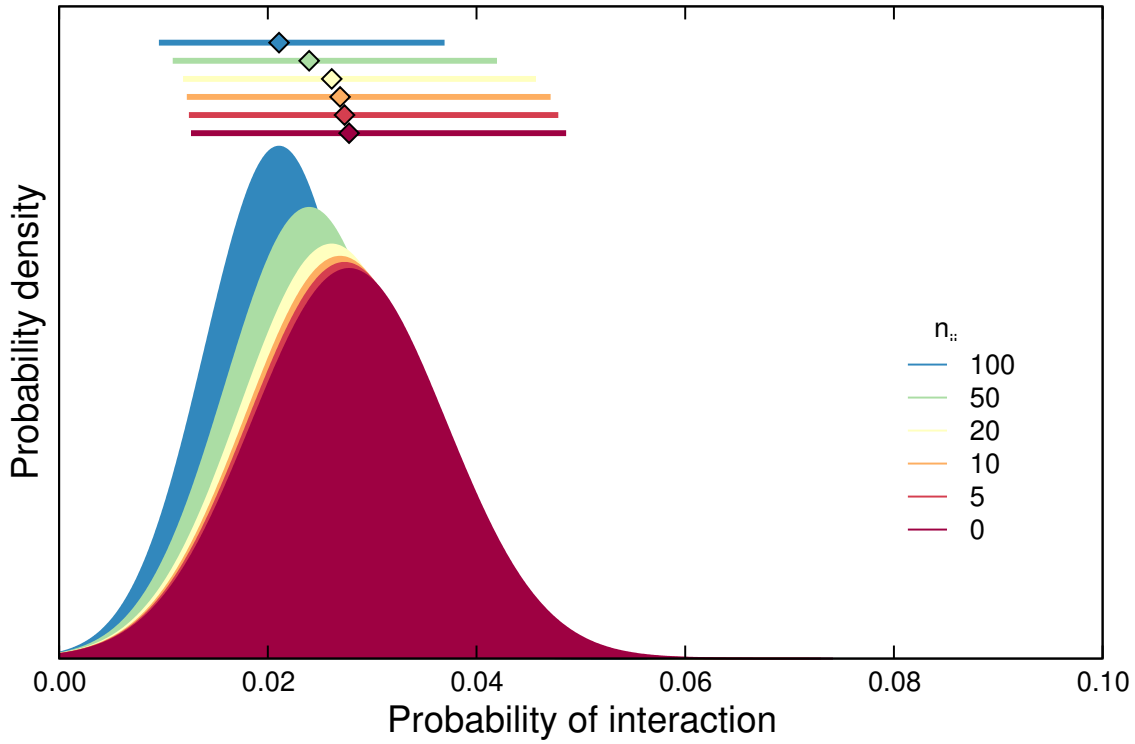
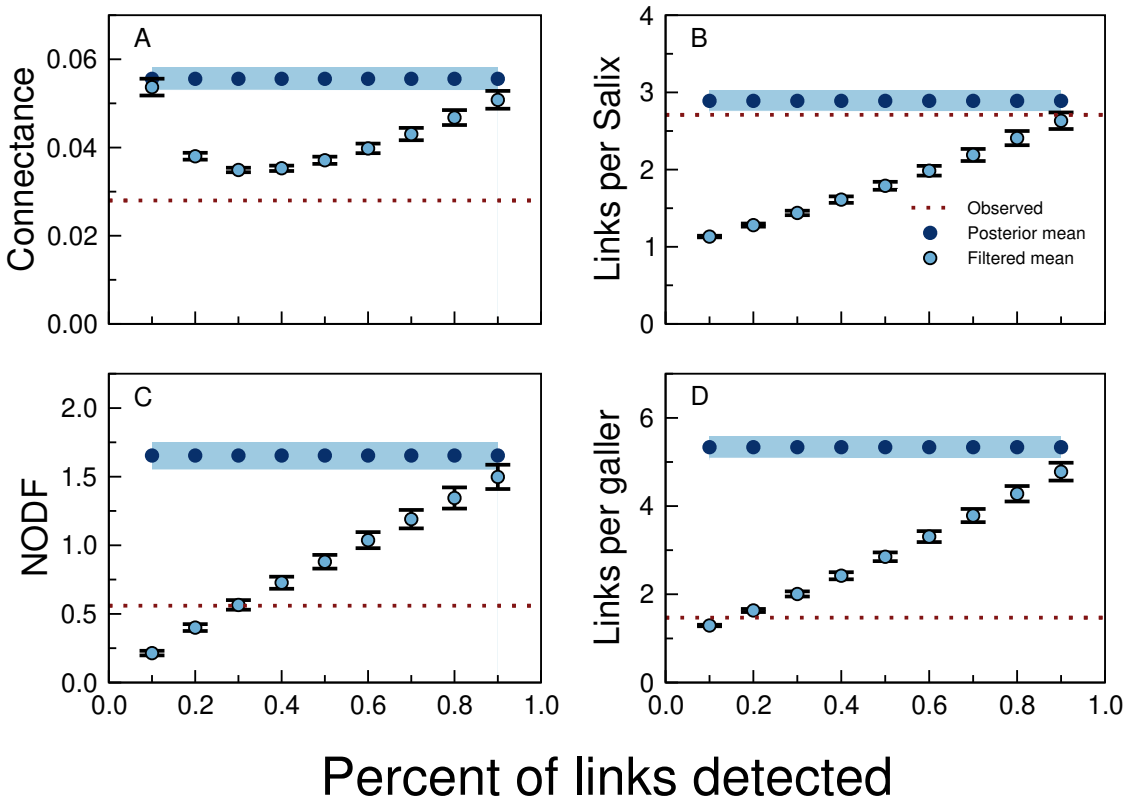


Figure S5: Here we show the mean connectance, links per resource (*Salix* in the *Salix*-galler networks and gallers in the galler-natural enemy networks), links per consumer, and nestedness (NODF) for networks assembled using posterior distributions based on a single sub-network in the Kopelke *et al.* (2017) dataset (Zillis). We created 100 "posterior-sampling" networks and then, for each of these, created 100 "detection-filter" networks by randomly sampling 50%-99% of the interactions included in the posterior-sampling network. This simulates imperfect detection of interactions in the field. Each point represents the mean network property (e.g., connectance) obtained from a set of 100 detection-filter networks, plotted against the value of the network property in the posterior-sampling network used to create the detection-filter networks. For each property and both network types, the posterior-sampling networks cover a relatively small range of network properties than the range covered by networks with varying detection probabilities. The value of each property decreases with the proportion of links included in the detection-filter networks.



Appendix S7: Analyses using alternative prior

We repeated our analyses using an alternative prior derived from a study of a similar *Salix*-galler-natural enemy system in North America (Barbour *et al.*, 2016; Barbour, 2016). We note that this study used several genotypes of *Salix hookeriana* rather than different *Salix* species and so did not produce networks of similar size and connectedness to those in (Kopelke *et al.*, 2017). To obtain the priors based on Barbour *et al.* (2016), we estimated frequencies of *S. hookeriana* genotype-galler interactions based on the normalised degree of each node (species or genotype) in each network component.

Creating a binary interaction network

Barbour *et al.* (2016); Barbour (2016) sampled 145 branches from *Salix hookeriana* of 26 genotypes. They recorded galls made by four species of Cecidomyid midges on each branch. We transformed these data to a binary genotype-galler network matrix, where an entry ij was 1 if galler i was observed on any branch of genotype j and 0 otherwise. In total, this network contained 75 realised galler-genotype interactions out of a potential 104. One *S. hookeriana* genotype did not interact with any gallers and was removed from the network. The R code used to extract this network follows:

```
prior_web_data=read.csv("tree_level_interaxn_all_plants_traits_size.csv")
# Remove extra data on parasitoids, etc.
prior_web_data[,31:57]<-NULL
prior_web_data[,4:26]<-NULL
prior_web_data[,2]<-NULL
# Build a web with interactions=1 for any galler observed on any genotype
prior_web=matrix(nrow=26,ncol=4)
for(r in 1:length(levels(prior_web_data$Genotype))){
    gen=levels(prior_web_data$Genotype)[r]
    subset=prior_web_data[which(prior_web_data$Genotype==gen),]
    for(c in 1:4){
     if(sum(subset[,2+c])>0){
      prior_web[r,c]=1
     } else {
      prior_web[r,c]=0
     }}}
# Remove non-interactive genotype to obtain the final web
prior_web<-prior_web[which(rowSums(prior_web)>0),]
```

To extract the galler-parasitoid network from the same study above:

```
prior_web_data=read.csv("tree_level_interaxn_all_plants_traits_size.csv")
prior_web_data <- prior_web_data[,8:26]
# Build a web with interactions=1 for any parasitoid on any galler
prior_web=data.frame()
# each column contains all interactions for a  particular parasitoid-galler combination
for(column in 1:ncol(prior_web_data)){
    # the column name gives the species which makes the gall,
    #followed by the species emerging from the gall
    galler_emerged=strsplit(colnames(prior_web_data)[column], "_")
    galler = galler_emerged[[1]][1]
    emerged = galler_emerged[[1]][2]
    # the strength of the interaction is the sum of the column
    (number of galls across all branches sampled)
    linkstrength = sum(prior_web_data[,column])
    prior_web[as.character(galler),as.character(emerged)] <- linkstrength
   }
# Make the NAs into 0
prior_web[is.na(prior_web)] <- 0
# remove emerging gallers with no parasitoids
prior_web_para_only <- prior_web[,c(2:6,8:9,12)]
# convert interaction strengths to 1
prior_web[which(prior_web>0)] <- 1
write.csv(prior_web_para_only, file="prior_web_para_only.csv", quote=F)
```

Prior and posterior distributions

Using the binary networks described above, we obtained prior parameters α=2.51 and β=1.89 for the *Salix*-galler component and α=1.34, β=9.49 for the galler-natural enemy component of the network. We then calculated prior distributions using these parameters. For the *Salix*-galler network, the prior distribution was: $\bar{\lambda}_{ij}$=0.570, var(λ_{ij})=0.056. The prior distribution for the galler-natural enemy network was: $\bar{\lambda}_{ij}$=0.124, var(λ_{ij})=0.010. As with the results in the main text, we can update this prior distribution using observed co-occurrences. In the most extreme case where a pair of species co-occurred at all 374 sites and was never observed interacting, our distributions would become $\bar{\lambda}_{ij}=6.63 \times 10^{-3}$, var($\lambda_{ij}$)=$1.74 \times 10^{-5}$ for the *Salix*-galler network and $\bar{\lambda}_{ij}=3.49 \times 10^{-3}$, var($\lambda_{ij}$)=$9.03 \times 10^{-6}$ for the galler-natural enemy network.

Credible intervals and sampling requirements

The 95% credible intervals calculated using the priors derived from Barbour *et al.* (2016) were much wider than those reported in the main text (Fig. S6). For hypothetical *Salix*-galler and galler-natural enemy pairs these intervals widened from (0.001, 0.017) and (<0.001, 0.11) for pairs observed co-occurring at all 374 sites without any observed interaction to (0.152,0.931) and (0.008, 0.364) for *Salix*-galler and galler-natural enemy pairs that were never observed co-occurring.

If we want to be 95% confident that the interaction probability for two species in the *Salix*-galler network is below 0.1, 0.05, or 0.01, we would need 51, 106, and 550 observed co-occurrences with no observed interaction, respectively (Fig. S7). The number of samples required to be 95% confident that the interaction probability between galler and natural enemy species is below a threshold also increases quickly as the threshold decreases; we would need 25, 62, and 352 observed co-occurrence with no observed interaction for threshold interaction probabilities of 0.1, 0.05, and 0.01, respectively.

Scaling up to network metrics

We find, perhaps not surprisingly, that the posterior webs for the Salix-galler network all had much higher connectance than the original, observed web (C=0.028 for the observed web and $0.528 \leq C \leq 0.568$ for the posterior webs; Fig. S8A). Likewise, the number of links per *Salix* and galler species in the observed web (L_{Salix}=2.71 and L_{galler}=1.47, respectively) were much lower than those predicted in the posterior webs ($27.4 \leq L_{Salix} \leq 29.5$ and $50.6 \leq L_{galler} \leq 54.5$, respectively; Fig. S8C,E), even when filtering with a detection probability of only 50%. Nestedness was also much higher in the posterior networks ($14.2 \leq NODF \leq 15.4$; Fig. S8G) than in the observed network (NODF=0.560). There are two possible explanations for these discrepancies: either the data of Barbour *et al.* (2016) is simply too different from that of Kopelke *et al.* (2017) to offer an appropriate prior for our dataset, or the true detection probability for links between *Salix* and galler species is much less than 50%. As the scale of the two datasets is quite different (genotypes of a single *Salix* species in Barbour *et al.* (2016), various *Salix* species in our dataset), we suspect the former is more likely.

Considering the galler-natural enemy networks, the connectance, mean links per galler, and mean links per natural enemy were also much lower in the observed web (C=0.078, L_{galler}=9.99, and $L_{naturalenemy}$=7.45, respectively) than in the posterior webs ($0.183 \leq C \leq 0.196$, $17.5 \leq L_{galler} \leq 18.9$, and $23.0 \leq L_{naturalenemy} \leq 24.8$). When the detection probability was relatively low (i.e., 50%), however, the properties of randomised networks became similar to those in the observed webs (Fig. S8B,D,F). Nestedness was higher in the observed network (NODF=6.85) than in the posterior networks ($6.30 \leq NODF \leq 6.74$; Fig. S8H), but network structure was much more similar in the observed and posterior galler-natural enemy networks than in the *Salix*-galler networks. The Barbour *et al.* (2016) network likely provides a better prior here than for the *Salix*-galler networks, as both networks included different galler and natural enemy species. Although the two networks differ vastly in scale, and the inclusion of multiple *Salix* species likely causes structural differences from Barbour *et al.* (2016) despite the similar resolution of gallers and natural enemies, this result nevertheless suggests that even such a large and well-replicated network as that in Kopelke *et al.* (2017) is missing many interactions. This should be a strong warning to researchers comparing structural characteristics between networks - small structural differences are likely to be masked by substantial noise resulting from sampling uncertainty.

Figure S6: Using prior distribution based on the *Salix*-galler and galler-natural enemy networks in Barbour *et al.* (2016), we can calculate posterior distributions for the probability of interaction (λ_{ij}) between two species that have not yet been observed interacting. Here we show posterior distributions for each λ_{ij} in each network ranging from the prior distribution ($n_{ij} = 0$ observed co-occurrence) to the distribution obtained when the pair of species has been observed co-occurring 100 times. The distribution narrows and approaches zero as the sample size increases. Likewise, the maximum likelihood estimator for the mean probability of interaction (diamonds at top of each panel) approaches zero and the 95% credible interval (lines at top of each panel) narrows as sample size increases. **A)** The posterior distributions for the *Salix*-galler component are always wider and farther from zero than those for **B)** the galler-natural enemy component. This is likely because the prior distribution for the galler-natural enemy component of the network was both narrower and had a mean closer to that in the Kopelke *et al.* (2017) data than did the *Salix*-galler component.

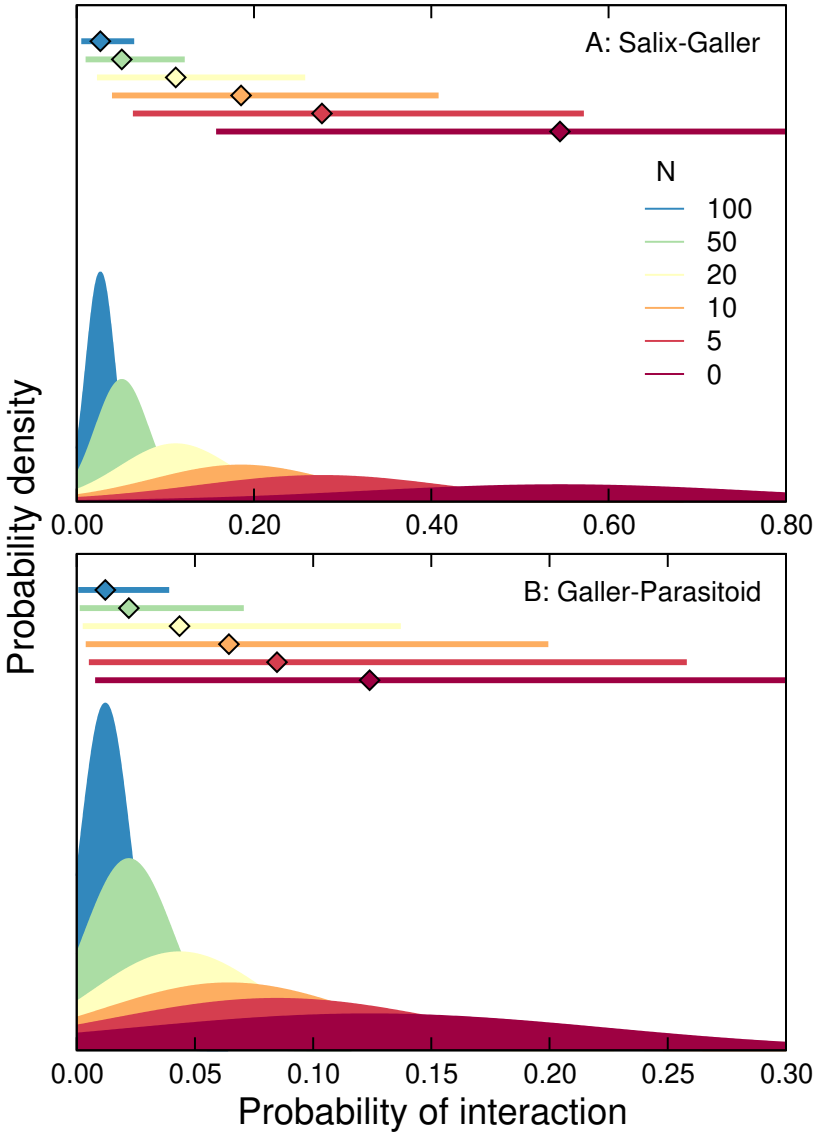


Figure S7: The number of samples required to achieve a given level of confidence that an interaction probability λ_{ij} is below a given threshold varies with both parameters. With a low threshold, our confidence that λ_{ij} is below the threshold increases rapidly with repeated observation of co-occurrence without interaction. Here we show the cumulative density functions for threshold probabilities of 0.5 (solid line), 0.25 (dashed line), 0.1 (dash-dot line), and 0.05 (dotted line) as well as the points at which the cdf reaches 0.90 (orange square), 0.95 (red circle), and 0.975 (blue diamond) for each threshold value. The large ticks along the x-axis indicate the number of samples associated with each of these points. Note that the number of samples required to reach any given threshold is larger for **A)** the *Salix*-galler network than for **B)** the galler-natural enemy network. In the galler-natural enemy network, the credible interval for a pair of species with no observed co-occurrences was 0.364. All pairs of species therefore have $\lambda_{ij}<0.5$.

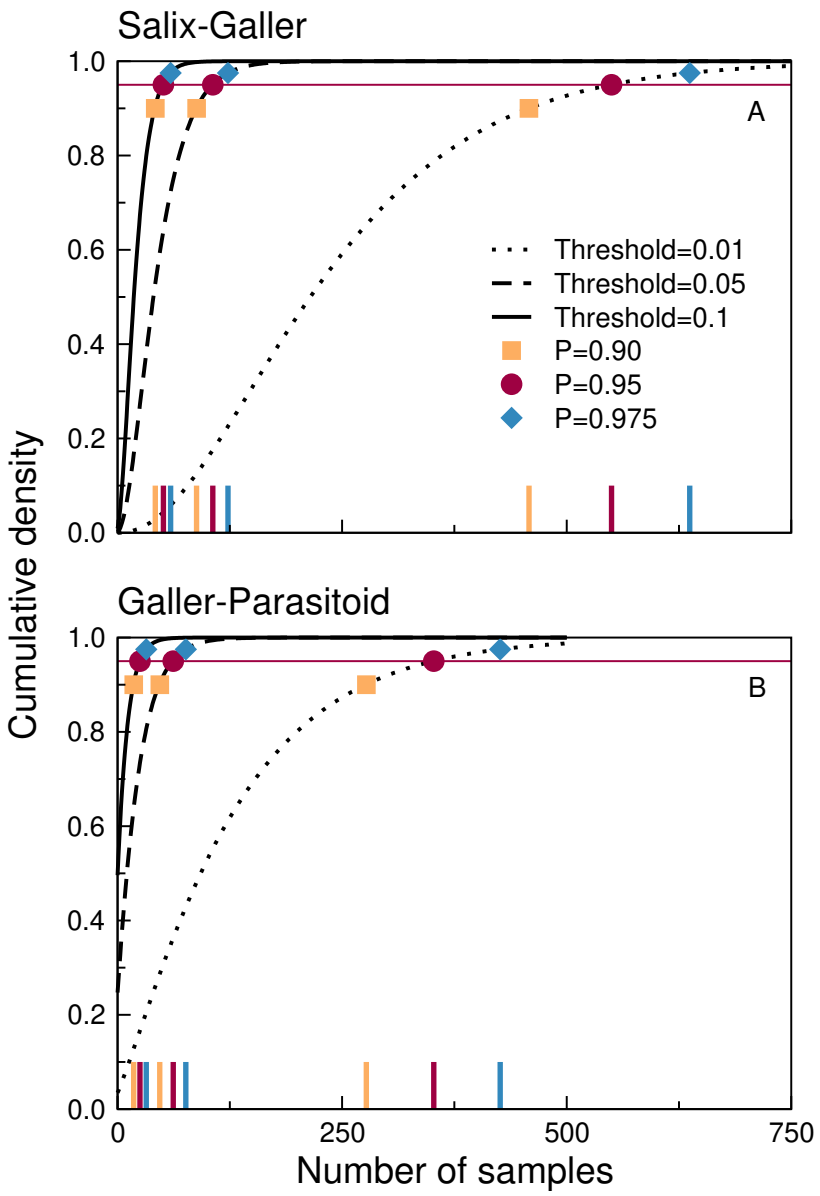
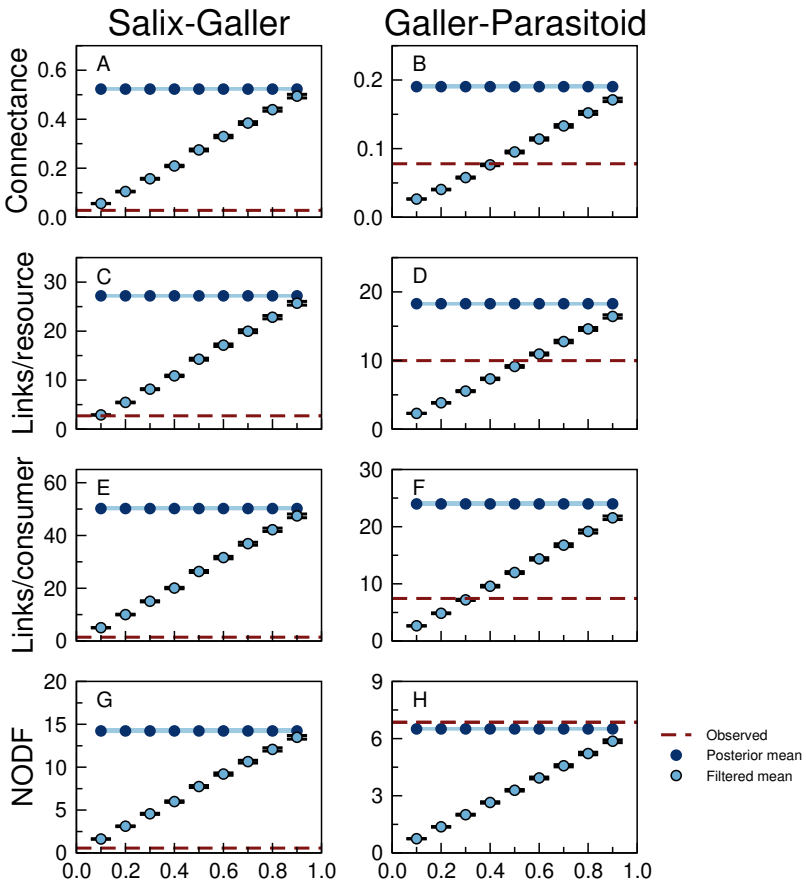


Figure S8: Here we show the mean connectance, links per resource (*Salix* in the *Salix*-galler networks and gallers in the galler-natural enemy networks), links per consumer, and nestedness (NODF) for networks assembled using posterior distributions based on a smaller, North American *Salix*-galler-natural enemy system (Barbour *et al.*, 2016; Barbour, 2016). We created 100 "posterior-sampling" networks and then, for each of these, created 100 "detection-filter" networks by randomly sampling 50%-99% of the interactions included in the posterior-sampling network. This simulates imperfect detection of interactions in the field. Each point represents the mean network property (e.g., connectance) obtained from a set of 100 detection-filter networks, plotted against the value of the network property in the posterior-sampling network used to create the detection-filter networks. For each property and both network types, the posterior-sampling networks cover a relatively small range of network properties than the range covered by networks with varying detection probabilities. The value of each property decreases with the proportion of links included in the detection-filter networks.



Appendix S8: Obtaining prior interaction probabilities

Degree distributions and interaction probabilities

Obtaining degree distributions from the binary interaction network is straightforward. To do this, we simply divide the total number of observed interaction partners for each species (row or column sums) with the number of potential interaction partners (number of column or rows). Degree distributions calculated, we can then calculate the probability of each interaction as the product of the normalised degrees of the two species. For example, the probability of an interaction between two species with normalised degrees of 0.5 (each interacts with half of the available partners) is 0.25. R code used to do this follows:

```
# Now get the degree distributions from the web
deg_dist_Salix=rowSums(prior_web)/ncol(prior_web)
# Remove one genotype that never interacted
deg_dist_Salix<-deg_dist_Salix[which(rowSums(prior_web)>0)]
deg_dist_galler=colSums(prior_web)/nrow(prior_web)
# Interaction probabilities are the product of plant and galler probabilities
int_probs=as.numeric(deg_dist_galler%*%t(deg_dist_Salix))
# And for the galler-parasitoid web:
deg_dist_parasitoid=rowSums(prior_web_para_only)/ncol(prior_web_para_only)
deg_dist_parasitised=colSums(prior_web_para_only)/nrow(prior_web_para_only)
gp_int_probs=as.numeric(deg_dist_parasitised%*%t(deg_dist_parasitoid))
```

Appendix S9: R code for calculating posterior distributions and confidence intervals

The following simple functions are implemented in the R language. They are also provided as R code in a separate file. Priordata is the list of interaction frequencies from the prior data, n_{ij} is the number of sites with observed co-occurrence of species i and j, and k_{ij} is the number of sites with an observed interaction ij. When calculating the prior distribution, both n_{ij} and k_{ij} are 0. In the main text we assume 100% confidence in observed interactions and therefore consider only cases where $k_{ij} = 0$. Only n_{ij} varies between species pairs.

To calculate the parameters α and β of a prior or posterior distribution:

```
calculate_parameters<-function(priordata,n,k){
    # Calculate prior parameters
    start=list(shape1=1,shape2=1)
    pars=fitdistr(x=priordata,'beta',start=start,lower=c(0,0))$estimate
    alpha=pars[[1]]
    beta=pars[[2]]
    # Update the parameters with data. If n=0 and k=0, no change.
    alpha_prime=alpha+k
    beta_prime=beta+n-k
    pars2=c(alpha_prime,beta_prime)
    return(pars2) }
```

All of the following functions use the parameters returned by "calculate_parameters". To calculate the maximum likelihood estimates of the mean and variance of the probability of interaction λ_{ij}:

```
calculate_distribution<-function(pars){
    # alpha and beta may be from a prior or posterior distribution
    alpha=pars[[1]]
    beta=pars[[2]]
    # Calculate the MLE of the mean
    mu_numerator=alpha
    mu_denominator=alpha+beta
    mu=mu_numerator/mu_denominator
    # Calculate the MLE of the variance
    sig_numerator=alphabeta
    den1=alpha+beta
    den2=den1**2
    sig_denominator=den1*den2
    sigma2=sig_numerator/sig_denominator
    return(c(mu,sigma2))}
```

To calculate a credible interval based on the prior or posterior distribution, for given lower and upper bounds:

```
credible_interval<-function(pars,p_lower,p_upper){
    alpha=pars[[1]]
    beta=pars[[2]]
    lowCI=qbeta(p=p_lower,shape1=alpha,shape2=beta)
    highCI=qbeta(p=p_upper,shape1=alpha,shape2=beta)
    return(c(lowCI,highCI))}
```

To calculate the number of samples required to reach a given level of confidence that the probability of interaction between two species (λ_{ij}) that have not been observed co-occurring is below a threshold value:

```
samples_for_threshold<-function(threshold,confidence,pars){
    alpha=pars[[1]]
    beta=pars[[2]]
    n=seq(0,100,1)
    k=0
    cdf=pbeta(threshold,shape1=alpha,shape2=beta+n)
    samples=length(which(cdf<confidence))
    return(samples) }
```

Appendix S10: Calculating network properties using posterior distributions

Once we have calculated posterior distributions for the probability of each interaction, we can simulate networks that might be observed if we could sample sufficiently to elminate some of the uncertainty in empirical networks. Using these posterior networks, it is straightforward to obtain a plausible distribution for network properties of the predicted "true" community. This extends earlier work establishing probabilistic forms of network properties (e.g., Poisot *et al.* (2015)) by allowing us to calculate higher moments of network properties rather than only the mean. Note that these network properties may differ substantially from those of the observed network. Here we demonstrate this process with the galler-parasitoid component of the network described in Kopelke *et al.* (2017).

Methods

Scaling up to network metrics

Researchers are often interested in measures of network structure rather than the network itself. Computing most network metrics is straightforward when the different λ_{ij} of the adjacency matrix are known and assumed not to vary (Poisot *et al.*, 2016). Incorporating variance in λ_{ij} into these calculations, however, is not so easy. Computation of these metrics involves non-linear functions. By Jensen's inequality (Jensen, 1906), this means that the average of a network metric (a function of stochastic interactions) is not the same as the network metric calculated based on the average interaction probability. Thus, any uncertainty in the values of λ_{ij} could bias both the mean and variance of network metrics, giving misleading results (simulated in *Appendix S1*). One way to avoid this situation is to calculate the properties of a suite of simulated networks (Vázquez *et al.*, 2005; Guimerà & Sales-Pardo, 2009).

Using the prior distributions and procedures described above, we calculated posterior probability distributions for species pairs that were not observed interacting. Using these posterior distributions and assuming probabilities of 1 for pairs of species that were observed interacting (see *Appendix S2* for a justification), we created a suite of 100 webs by randomly sampling from each posterior distribution. Each of these webs is a prediction of the structure of the metaweb. After obtaining these posterior networks, we calculated, as examples of commonly-used network properties, the connectance of each web, as well as the mean number of links per galler and per natural enemy, and the nestedness (NODF) of the network. This gives us an estimate for each property reflecting our uncertainty about λ_{ij}.

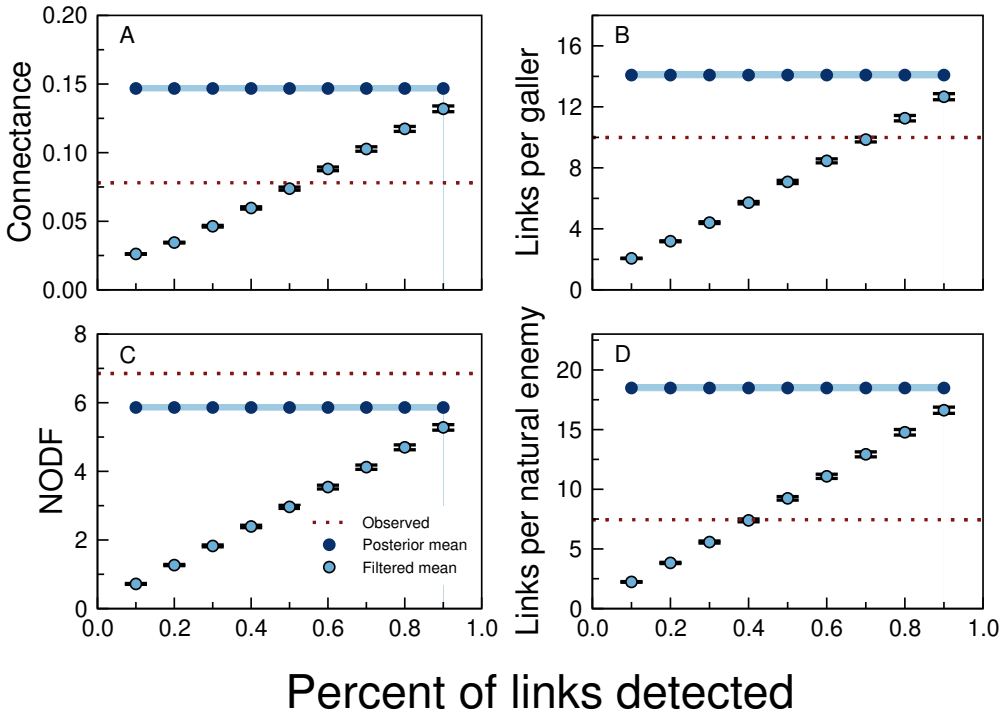
Measures of network structure which are based on empirical networks that are missing interactions may differ substantially from the values that would be obtained if detection certainty and variation in interactions over space and time could be removed. To demonstrate this, we created a suite of filtered networks for each posterior network. Taking a posterior network as the "true" network, we randomly sampled 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, and 10% of the interactions to create a new "filtered" network. This gradient is akin to a gradient of uncertainty in the network (e.g., due to varying sampling effort or intraspecific trait variability). For each level of uncertainty, we created 100

randomly-sampled networks per posterior-probability network (giving 100 posterior networks and 1000 filtered networks). We calculated the same network properties as described above for all posterior and filtered networks.

Results

Scaling up to network structure, we found that the connectance and mean links per galler and natural enemy were much lower in the observed web (C=0.078, L_{galler}=9.99, and $L_{naturalenemy}$=7.45, respectively) than in the posterior webs ($0.186 \leq C \leq 0.198$, $13.4 \leq L_{galler} \leq 14.6$, and $23.4 \leq L_{naturalenemy} \leq 25.0$). When the detection probability was relatively low (i.e., 50%), however, the properties of randomised networks became similar to those in the observed webs (Fig. S9A,B,D). Nestedness was higher in the observed network (NODF=6.85) than in the posterior webs ($6.31 \leq NODF \leq 6.82$; Fig. S9C), indicating that the posterior webs include many more interactions between specialists than the observed network. In addition, the stronger the detection filter, the farther apart were the nestedness of the observed and posterior webs. This suggests that the interactions included in the observed network are not a random subset of those included in the posterior webs. In general, the observed network is most similar to simulated networks where only half of the plausible links are detected.

Figure S9: Mean connectance, links per galler, nestedness (NODF), and links per natural enemy for networks assembled using posterior distributions based on a single site (Zillis in Graubünden, Switzerland) from Kopelke *et al.* (2017) occupied a narrow range. To obtain distributions of network properties, we created 100 "posterior-sampling" networks and then, for each of these, created 100 "filtered" networks by randomly sampling 50%-99% of the interactions included in the posterior-sampling network. This simulates imperfect detection of interactions in the field. Each point represents the mean (+/- SD) network property (e.g., connectance) obtained from a set of 100 filtered networks. The filtered networks cover a broader range of network properties than the posterior-sampling networks but the value of each property decreases as the strength of the detection filter increases. Values in the original web are indicated by dashed lines.



Appendix S11: Calculating the credible interval around a probability estimate

Here we describe the derivation of the Clopper-Pearson credible interval for the estimated probability of interaction λ_{ij} of a pair of species observed co-occurring n_{ij} times and interacting k_{ij} times. As we are most interested in the probability of interaction between species pairs that have never been observed co-occurring, we consider only the case where $k_{ij} = 0$ over a variety of n_{ij}. This is straightforward to do in R (see the function "credible_interval" in *Appendix S7*).

First, we must obtain the α and β parameters for the prior distribution. In this study we obtained these parameters using the R (R Core Team, 2016) function fitdist from the package fitdistrplus (Delignette-Muller & Dutang, 2015). Once α and β are known, we can update them using our observed data. Specifically, we are interested in $\alpha' = \alpha + k$ and $\beta' = \beta + n - k$. These parameters can then be used to calculate a credible interval using the R (R Core Team, 2016) function qbeta. In the table below, we present the 95% credible intervals for *Salix*-galler and galler-natural enemy pairs with different numbers of observed co-occurrences (n_{ij}) and no observed interactions ($k_{ij} = 0$), calculated using prior information derived from the Zillis sub-network (Kopelke *et al.*, 2017).

Table S2: Here we give the lower and upper bounds of 95% credible intervals for the probability of interaction λ_{ij} between *Salix*-galler or galler-natural enemy pairs that have been observed co-occurring n_{ij} times but have never been observed interacting.

n_{ij}	*Salix*-galler		galler-natural enemy	
	Lower bound	Upper bound	Lower bound	Upper bound
0	0.013	0.049	5.39×10^{-4}	0.304
1	0.013	0.048	4.82×10^{-4}	0.276
2	0.013	0.048	4.35×10^{-4}	0.253
5	0.012	0.048	3.37×10^{-4}	0.203
10	0.012	0.047	2.45×10^{-4}	0.152
15	0.012	0.046	1.93×10^{-4}	0.121
20	0.012	0.046	1.59×10^{-4}	0.101
25	0.012	0.045	1.35×10^{-4}	0.087
50	0.011	0.042	7.72×10^{-5}	0.050
100	0.010	0.037	4.16×10^{-5}	0.027
150	0.009	0.033	2.84×10^{-5}	0.019
200	0.008	0.030	2.16×10^{-5}	0.014
374	0.006	0.022	1.18×10^{-5}	0.008

References

Barbour, M.A. (2016). Data from: *Genetic specificity of a plant-insect food web: implications for linking genetic variation to network complexity*. Dryad Digital Repository. Available at: https://doi.org/10.5061/dryad.g7805.

Barbour, M.A., Fortuna, M.A., Bascompte, J., Nicholson, J.R., Julkunen-Tiitto, R., Jules, E.S. & Crutsinger, G.M. (2016). Genetic specificity of a plant-insect food web: implications for linking genetic variation to network complexity. *Proceedings of the National Academy of Sciences of the United States of America*, 113, 2128–2133.

Bartomeus, I. (2013). Understanding linkage rules in plant-pollinator networks by using hierarchical models that incorporate pollinator detectability and plant traits. *PLoS ONE*, 8, e69200.

Brown, L.D., Cai, T.T. & DasGupta, A. (2001). Interval estimation for a binomial proportion. *Statistical Science*, 16, 101–133.

Carstensen, D.W., Sabatino, M., Trøjelsgaard, K. & Morellato, L.P.C. (2014). Beta diversity of plant-pollinator networks and spatial turnover of pairwise interactions. *PLoS ONE*, 9, e112903.

Delignette-Muller, M.L. & Dutang, C. (2015). fitdistrplus: an R package for fitting distributions. *J Stat Softw*, 64, 1–34.

Graham, C.H. & Weinstein, B.G. (2018). Towards a predictive model of species interaction beta diversity. *Ecology Letters*, 21, 1299–1310.

Gravel, D., Baiser, B., Dunne, J.A., Kopelke, J.P., Martinez, N.D., Nyman, T., Poisot, T., Stouffer, D.B., Tylianakis, J.M., Wood, S.A. & Roslin, T. (2018). Bringing Elton and Grinnell together: a quantitative framework to represent the biogeography of ecological interactions. *Ecography*, Online Early View.

Guimerà, R. & Sales-Pardo, M. (2009). Missing and spurious interactions and the reconstruction of complex networks. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 22073–22078.

Jensen, J.L.W.V. (1906). Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Math*, 30, 175–193.

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

Kaartinen, R. & Roslin, T. (2011). Shrinking by numbers: landscape context affects the species composition but not the quantitative structure of local food webs. *Journal of Animal Ecology*, 80, 622–631.

Kopelke, J.P., Nyman, T., Cazelles, K., Gravel, D., Vissault, S. & Roslin, T. (2017). Food-web structure of willow-galling sawflies and their natural enemies across Europe. *Ecology*, 98, 1730.

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.

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

R Core Team (2016). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Roslin, T. & Majaneva, S. (2016). The use of DNA barcodes in food web construction–terrestrial and aquatic ecologists unite! *Genome*, 59, 603–628.

Vázquez, D.P., Morris, W.F. & Jordano, P. (2005). Interaction frequency as a surrogate for the total effect of animal mutualists on plants. *Ecology Letters*, 8, 1088–1094.

Wirta, H.K., Hebert, P.D.N., Kaartinen, R., Prosser, S.W., Várkonyi, G. & Roslin, T. (2014). Complementary molecular information changes our perception of food web structure. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 1885–1890.