Breaking the spell of nestedness

Clàudia Payrató Borrás[1,2], Laura Hernández[1], and Yamir Moreno[2,3,4,5]

Abstract
Mutualistic interactions, which are beneficial for both interacting species, are recurrently present in ecosystems. Observations of natural systems showed that, if we draw mutualistic relationships as binary links between species, the resulting bipartite network of interactions displays a widespread particular ordering called nestedness [1]. On the other hand, theoretical works have shown that a nested structure has a positive impact on a number of relevant features ranging from species coexistence [2], to a higher structural stability of communities and biodiversity [3, 4]. However, how nestedness emerges and what are its determinants, are still open challenges that have led to multiple debates to date [5–7]. Here, we show, by applying a theoretical approach to the analysis of 167 real mutualistic networks, that nestedness is not an irreducible feature, but a consequence of the degree sequences of both guilds of the mutualistic network. Remarkably, we find that an outstanding majority of the analyzed networks does not show statistical significant nestedness. These findings point to the need of revising previous claims about the role of nestedness and might contribute to expand our understanding of how evolution shapes mutualistic interactions and communities by placing the focus on the local properties rather than on global quantities.

Keywords
Mutualistic Ecosystems — Nestedness – Exponential Random Graph – Null Models

[2] *Laboratoire de Physique Théorique et Modélisation, UMR CNRS, Université de Cergy-Pontoise, 2 Avenue Adolphe Chauvin, F-95302, Cergy-Pontoise Cedex, France.*
[2] *Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Spain.*
[3] *Department of Theoretical Physics, Faculty of Sciences, University of Zaragoza, Spain.*
[4] *Complex Networks and Systems Lagrange Lab, Institute for Scientific Interchange, Turin, Italy.*
[5] *Complexity Science Hub Vienna, Austria.*

Nested patterns are ubiquitous in ecological systems. This observation has triggered an intense research aimed at defining and measuring nestedness [8] as well as at explaining its origin [9–11]. The interest in deciding whether a system is nested or not goes beyond characterizing it from a merely topological viewpoint. Admittedly, it has been suggested that nestedness plays an important role in biodiversity persistence, a claim which is nevertheless the subject of an ongoing and intense debate [2–4]. Furthermore, the relevance of nestedness as a suitable indicator to characterize mutualistic ecosystems has been recently challenged [12, 13]. Several works have proposed alternative properties of the observed networks to link relevant structural characteristics with the system's dynamics, in particular, the networks' assortativity or the heterogeneity in the number of interactions of the species [13–15].

A key question is then whether nestedness, conceived as a global trait of the emerging architecture, is actually relevant and informative of the ecosystems' dynamics, or contrarily, it just derives from lower order properties of the interaction network. In addition, elucidating the previous question would also allow to solve another open challenge, namely, which is the right *null model* against which one should assess nestedness? The latter is a relevant issue by itself, as any claim concerning statistical significance of a nested pattern implicitly involves its comparison with a null hypothesis (model). In order to address the aforementioned questions, we analyze an empirical set of 167 mutualistic networks (see *Methods*) to determine if, indeed, the observed amount of nestedness in real ecosystems could solely arise from the empirical degree sequences. Our choice is rooted on a a theoretical work [16] that showed that the geometric curve that delimits the region with interactions in an ideally nested matrix [17] can be ultimately related, by means of an approximation, to the *degree distributions* of both guilds of the corresponding bipartite network.

We constructed a grand canonical ensemble for each empirical ecological web under the constraint that, for the two guilds, the degree sequences in the ensemble match on *average* the empirical ones (see *Methods* and the *SM*). This methodological approach has the advantage that possible missing links or overrated interactions, that have been suggested to lead to impoverished ecological data [18], are dealt with in a proper way. In fact, constraining the randomized degree sequences to be equivalent to the empirical ones only on average limits the possible effects of noisy data, while assuring that results are not dependent on specific details. At variance with previous works that have imposed similar constraining rules [19], here we apply a recently introduced randomizing scheme [20, 21] that treats the ensemble from a statistical physics perspective, yielding the maximum entropy network ensemble such that the degree sequence of the empirical network is found with maximum likehood (see *Methods*). This

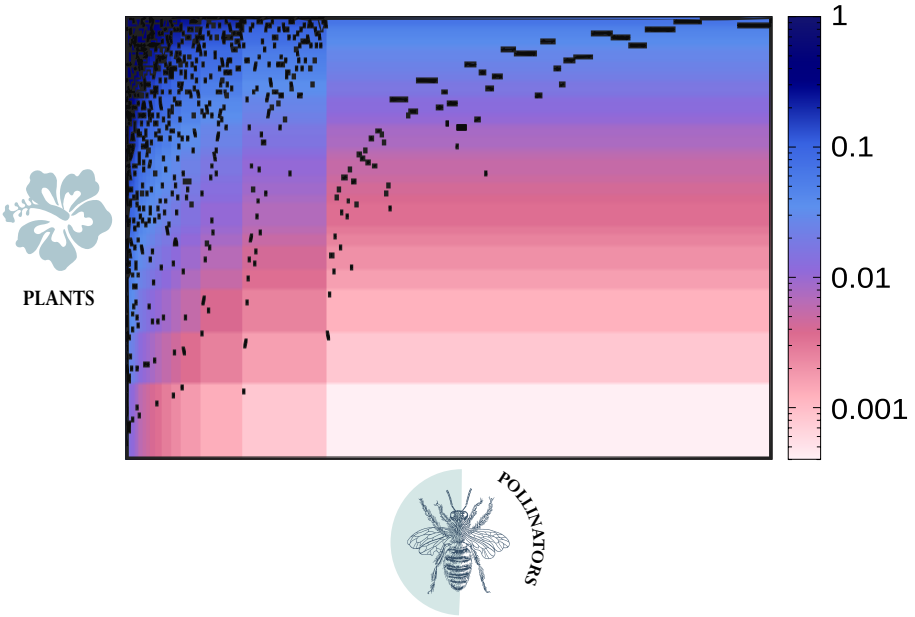

Figure 1. Comparison of generated and empirical mutualistic interactions. Probability of interaction between species in the statistical ensemble (color coded as indicated), for the plant-pollinator network recorded by Inoue et al. [23]. The empirical corresponding bipartite matrix of interactions is superimposed in black. Both plants and pollinators species have been ordered in decreasing order of their degrees (from top to bottom and from left to right). As it can be seen at a glance, the obtained probabilities are consistent with the observed interactions, with the dark regions delimiting an upper left triangle, as in an ideally nested structure. Note that the color legend is in logarithmic scale.

enforces, as aimed, the ensemble's mean degree sequences to be the empirical ones whilst precluding common biases of other sampling techniques [22], and allows us to obtain the probability that two potential partners interact in the randomized ensemble, see Fig.1. Moreover, one can explicitly write expressions for the main statistical moments of any network property that can be analytically formulated in terms of the elements of the bipartite adjacency matrix. Since the well-known NODF metric for nestedness fulfills such a condition, we have derived the analytical expressions for the mean and the standard deviation of the nestedness' distribution, see *Methods*.

For each one of the 167 empirical networks we have obtained the interaction probability between the elements of the corresponding bipartite matrix. These networks include three different kinds of mutualistic communities: plant-pollinator, seed-disperser and plant-ant (see Section SI5 of the *SM*). A comparison between the obtained mean expectations of the nestedness of the randomized ensembles and the measured values of the nestedness of the real networks unveils a striking agreement, see Fig. 2. As reported in Table 1, the absolute difference between these two quantities is less than *one* standard deviation for 100 out of 167 networks (59.9%), raising to 158 out of 167 networks (94.6%), if we account for *two* standard deviations. The previous percentage increases further after performing a multiple testing correction (see *Methods*): we find that only 3 out of the 167 empirical observations of nestedness are significant (p-value < 0.05). The three of them, which are of a relatively small size (≤ 55 species), were found to be *less* nested than predicted by the statistical ensemble. Additionally, in order to ensure that our findings are not an artifact of using the NODF metric, we have performed the same analysis by using the largest eigenvalue radius, which has been recently proposed as an alternative way of measuring nestedness [12]. In this case, since it is not possible to obtain an analytical and derivable expression of the metric, we broadly sampled the statistical ensemble using the calculated probabilities and performed the statistical measures on the obtained samples (see *Methods*). This supplementary analysis produces results that are in agreement with those reported above for the NODF, with only 16 out of 167 networks unexpectedly nested (see Section SI3 of the *SM*).

The findings above are of utmost importance in at least two fundamental aspects. Firstly, they demonstrate that, given the degree sequence of real networks, the observed nestedness is not significant. Secondly, they show that nestedness is not an irreducible pattern, in sharp contrast to the widely extended belief that it is so. In other words, these results reveal that the

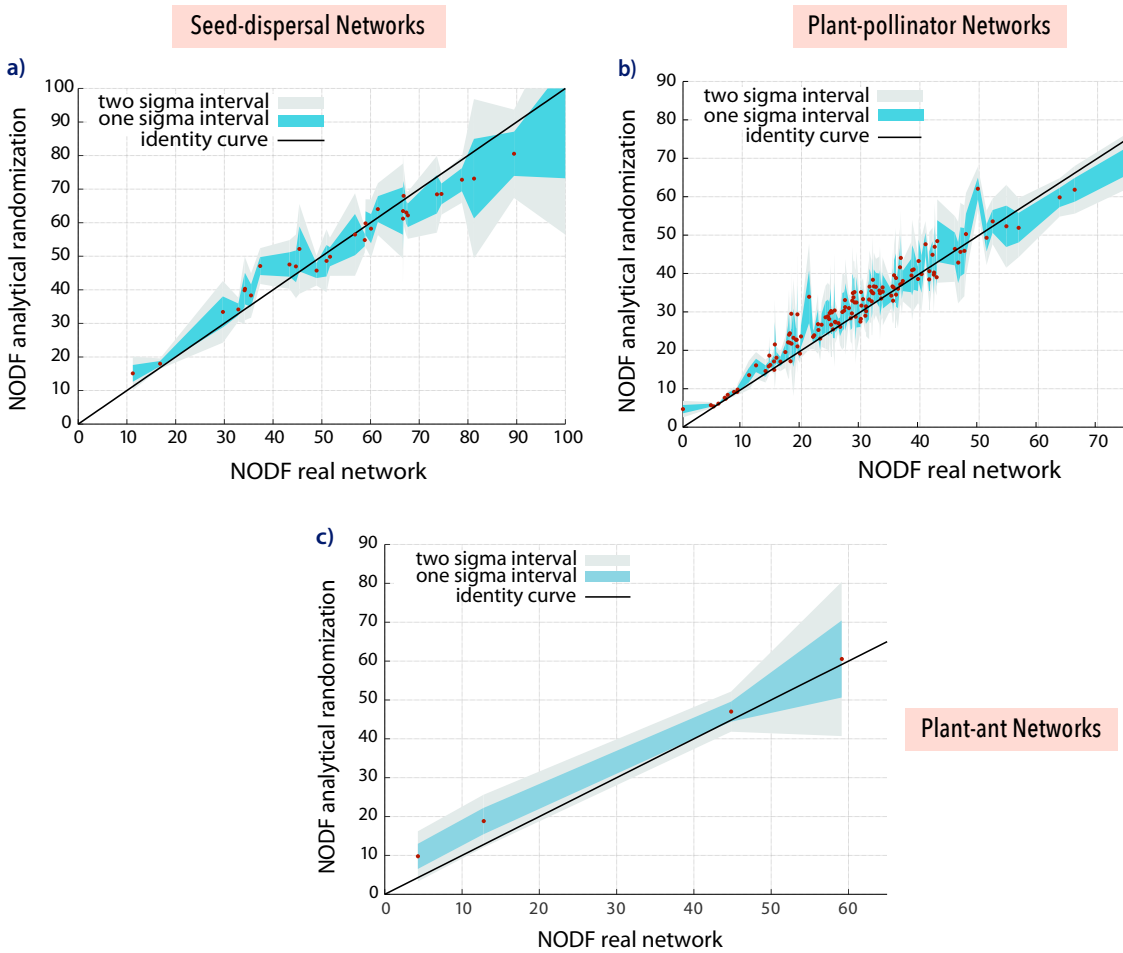

Figure 2. Determinants of nestedness. Panel *(a)*: relative change in nestedness and the corresponding change in heterogeneity, measured for the set of 167 empirical networks and the average over the respective rewired ones. We used the rewiring algorithm described in *Methods*. Nestedness is measured using the NODF metric, whereas the heterogeneity is measured through the variance of the degree sequence of the unipartite adjacency matrix. We found a correlation index for a linear fit (excluding the top outlier) of $R = 0.88$. This closely linear relationship discovers a tight bound between nestedness and heterogeneity. Panel *(b)* shows a comparison between the real observation of the degree assortativity r (Pearson's coefficient among degrees) and the average estimation in the statistical ensemble, for the 167 networks of our study. The fact that $r < 0$ for all values indicate that both real networks and the average of the randomized ensemble are naturally disassortative.

observed nested structure of the ecological communities studied is, in fact, a mere consequence of the degree sequences of the two guilds. Moreover, regarding recent debates about the use of a proper null model for nested networks [24], our findings point out the need of incorporating the information contained in the degree sequences. Indeed, our results indicate that an appropriate null model is the set of exponential random graphs for which the probability of having the same degree sequences of both branches of the bipartite graphs when compared to a real mutualistic network is maximized. Thus, we propose that the methodology implemented here to obtain the statistical ensemble of graphs that are compatible with the real networks could be a general tool to assess nestedness' significance.

In the light of the previous results, the second question arising is whether we can determine which characteristic of the

Type of community	fraction of ntws with |z-score| ≤ 1		fraction of ntws with |z-score| ≤ 2	
plant-pollinator	82 out of 133	61.7%	126 out of 133	95.5%
seed-disperser	16 out of 30	53.3%	28 out of 30	93.3%
plant-ant	2 out of 4	50.0%	4 out of 4	100.0%

Table 1. Results, disentangled into communities, showing the fraction of networks (abbreviated above as 'ntws') whose discrepancy between the real and randomized nestedness is less or equal than one or two standard deviations.

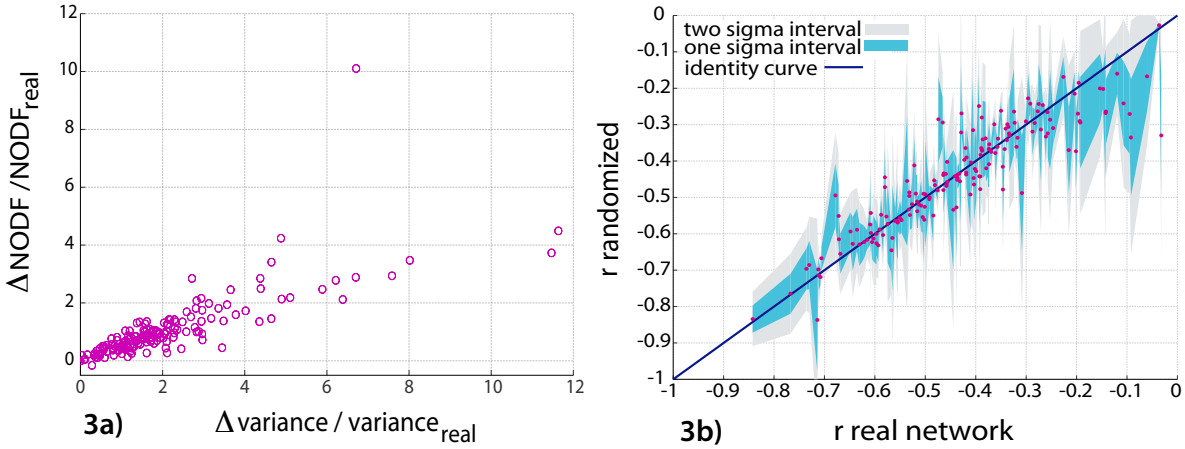

Figure 3. Determinants of nestedness. Panel *(a)*: relative change in nestedness and the corresponding change in heterogeneity, measured for the set of 167 empirical networks and the average over the respective rewired ones. We used the rewiring algorithm described in *Methods*. Nestedness is measured using the NODF metric, whereas the heterogeneity is measured through the variance of the degree sequence of the unipartite adjacency matrix. We found a correlation index for a linear fit (excluding the top outlier) of $R = 0.88$. This closely linear relationship discovers a tight bound between nestedness and heterogeneity. Panel *(b)* shows a comparison between the real observation of the degree assortativity r (Pearson's coefficient among degrees) and the average estimation in the statistical ensemble, for the 167 networks of our study. The fact that $r < 0$ for all values indicate that both real networks and the average of the randomized ensemble are naturally disassortative.

degree sequences modulates how nested a network is. Considering that the degree distributions of mutualistic communities have been reported to commonly follow a (truncated) power-law [25], we propose, as a plausible candidate, the *heterogeneity* in the number of contacts per species. Thus, our hypothesis is that for two networks with identical number of species and connections but diverse degree sequences, the most heterogeneous one (taking into account both guilds) will be as well the most nested. To evaluate this conjecture, we made use of a self-organizing network model that is devised with the aim of optimizing the nestedness of a network [5] by rewiring existing links (see *Methods*). After applying this algorithm to our empirical set of networks, we found that the resulting degree sequences are, with respect to the original ones, more heterogeneous and that the final networks are more nested, see Fig. 3a. This allows to bridge the gap between two topological features that have been classically treated separately, though previous works already suggested their connection [11, 14]. Interestingly enough, the relationship between network's heterogeneity and nestedness also explains why dynamical implications once attributed to nestedness like the sustainability of communities with a large number of different coexisting species [2] or the network's structural stability [3, 26], have recently been successfully associated with other properties such as the heterogeneity itself [15] or the species' degree [13].

Moreover, accounting for the heterogeneity offers some further insight on the process of emergence of nestedness out of the degree sequences. At first glance, it might not be evident why our null model reproduces so well the empirical nestedness. A priori, we would naively expect that the random ensemble contains both nested and non-nested structures alike, in which specialists appear attached, respectively, to generalists or to other specialists. Although a given number of connections are certainly imposed by the existence of super-generalists as well as by finite size effects, normally there is still room for reshuffling links (like in the "swapping algorithm" [27]). In terms of mixing, we would say that, concerning specialists, both assortative configurations (nodes have neighbors with degrees similar to their own) and disassortative ones (neighbors have dissimilar degree) are in theory feasible. Why, then, our algorithm is expected to generate disassortative networks (see Fig. 3b)? Here, the particularity that we used a maximally-entropic ensemble plays a crucial role. Johnson et al. [28] showed that, in the case of heterogeneous systems, disassortativity is generally more entropic, that is, it is more likely as long as no external pressures are at work. This occurs, to put it simply, because for a species with few interactions there exist many more chances to engage with another species with numerous connections than matching to a low-connected partner. Therefore, the low significance of empirical nested patterns reported here is directly related to the fact that the number of mutualistic interactions per species is a highly heterogeneous quantity.

In concluding, it is worth mentioning that in recent years, nestedness has been proposed to arise either as an ecological feature that provides an optimal balance between competition and mutualism [2, 6], or as a byproduct of processes such as the assembling rules [7, 29]. Our results imply that no selective pressure has acted upon nestedness, which does not exclude, however, that such pressure has shaped the degree sequences. Even though such conclusions do not invalidate nestedness'

usefulness as an indicator of stability or robustness, we would like to underline that our findings clearly demonstrate that the degree sequences are the lower-order determinants of nestedness. Moreover, this highlights the interest of focusing on the ecological and evolutionary mechanisms that have led to the heterogeneous degree distributions present in mutualistic ecosystems [**?**, **?**], like might be the need to diminish the cost of mutualism [**?**]. Understanding the way in which structural properties emerge in ecological communities is a fundamental, long-standing challenge that can provide critical clues to depict ecosystems' past assembling, present functioning and future responses. Finally, given that nested patterns have been recurrently detected across systems as diverse as biological, social and technological networks, our findings are expected to have relevant implications beyond the present analysis of ecological mutualistic communities.

Methods

Construction of the Random Ensemble. We constructed an ensemble following the Exponential Random Graph model. This ensemble maximizes the Shannon-Gibbs Entropy given the average degree sequences of the two guilds of a bipartite network as constraint. Yet, it is not fully determined due to the presence of some free Lagrange multipliers resulting from the constrained optimization. Following Squartini and Garlaschelli [20] [22], we imposed that the degree sequences of the empirical network are found with maximum likelihood. This provides a set of coupled equations to solve for the Lagrange multipliers (one equation per node, see Eqs. 7-8 in Section SI1 of the *SM*). Determining the statistical random ensembles of our 167 empirical networks entails solving computationally 167 optimization problems. For each network, we numerically found the Lagrange multipliers that maximize the likelihood using two different, independent algorithms: *1)* a global, pseudo-random numerical method for optimizing the likelihood and *2)* a deterministic, gradient-based algorithm for solving non-linear systems of equations. See Section SI1 of the *SM* for the explicit expressions of the ensemble probability and equations to solve, as well as more information on the numerical implementation. As shown in [20], in the case of local constraints (as the degree sequences), the probability of existence of a graph in the ensemble can be exactly factorized into the probabilities of existence of a link between species [20]. Therefore, after numerically determining each optimal set of Lagrange multipliers, we built the matrix containing the average probability of interaction corresponding to each empirical network (see an example in Fig. 1 and Eqs. 9-10 in Section SI1 of the *SM*).

Statistical Measures on the Random Ensemble. We performed the statistical measures on the ensemble following either of the two following approaches. On the one hand, as long as the property that we aim to evaluate could be formulated as an analytical and derivable expression, Squartini and Garlaschelli showed [20] that it is possible to obtain, at first order, the analytical expression of the first and second moments of the corresponding distribution. These expressions depend only on the link probabilities (see Eq. 1-2 in Section SI2 of the *SM*). We wrote the NODF metric in a compact, analytical form and derived the expression of the theoretical average expectation and standard deviation of nestedness in the ensemble (Eqs. 7-9 in Section SI2 of the *SM*). Finally, we obtained the 167 probability matrices of interactions and computed the main statistical moments. On the other hand, one can always sample the ensemble in order to study the statistics of the target property on a generated sampling. Using this scheme, we produced 10^4 networks that were assembled using the obtained probability matrix of interactions. Over this subset, we numerically calculated the average expectation and the standard deviation of the largest eigenvalue radius [12] (see Section SI3 of the *SM*) and the assortativity index measured through the Pearson coefficient of the degrees (see Section SI4 of the *SM*).

Significance tests. We quantified the significance of the nestedness using the *z-score* index, which for a general property x reads: $\frac{x^* - \langle x \rangle}{\sigma_x}$. For us, $\langle x \rangle$ is the average nestedness computed in the ensemble, and we compare it with the empirical observations x^*. The standard deviation is σ_x. Given that the NODF values are gaussian distributed in the random ensemble (see Section SI2 of the *SM*), the z-scores can be directly related to p-values. We performed a *multiple test correction* which allows accounting for the fact that as the number of statistical tests increases, so does the probability of finding rare events [30]. Thus, when considering the multiple comparisons we could prevent overstating the number of significant discoveries. It is pertinent to apply this technique here since the 167 cases studied are evaluated under the same *null hypothesis* and all of them follow a normal distribution. We employed the *false discovery rate* method, in particular the Benjamini-Hochberg procedure which applies to independent tests [30].

Self organizing network model. In order to reorganize the original network into an even more nested structure, we numerically implemented the self-organizing network model proposed by Burgos et al. [5]. This methodology keeps constant many aspects susceptible to affect the measure of nestedness, like the size and fill, but modifies the degree sequences through the redistribution of connections. We rewired the links among species following two simple rules: *i)* when changing an interaction, the new partner must have higher degree than the old neighbor *ii)* if the proposed redistribution leaves one of the two nodes with no interactions at all, we reject the change. This operation was repeated until the system achieved a frozen state in which no more reconnections were accepted (we considered this happened when $10^3 N$ consecutive rejections occurred, being N the

number of nodes of the network). The final frozen state is normally not perfectly nested, since condition *ii* typically leads to configurations which are not utterly optimal. To compensate this, we carried out 10^3 independent rewiring operations for each network. We then averaged the target properties, namely, nestedness (measured using NODF) and the variance of the joint degree sequence of the two guilds.

SUPPLEMENTARY INFORMATION

S1. Construction of the random ensemble

This Section provides additional details on how we constructed the statistical ensembles. An ensemble is a set of networks across which unconstrained features will vary randomly, and over which we will perform statistical measures.

General randomizing scheme

We denote a network in the ensemble by its graph $\mathbf{G}$, except for the real network which is $\mathbf{G}^*$. We characterize the ensemble by the probability of occurrence of each of its elements, $P(\mathbf{G})$. Following [31] and [20], we choose a probability such that the constraints are only satisfied on average, thus allowing slight mismatches across the ensemble. This is equivalent to constructing a *grand-canonical ensemble* (as opposite to the *micro-canonical ensemble*, where the constraints need to be exactly met always).

As proposed in [31], we ask the probability of each graph in the ensemble to maximize the Shannon-Gibbs entropy, defined as:

$$S = -\sum_{\mathbf{G}} P(\mathbf{G}) lnP(\mathbf{G}), \tag{1}$$

where the sum runs over all the graphs G in the ensemble. This leads to the *Exponential Random Graph model*, which reads:

$$P(G/\vec{\theta}) = \frac{e^{-H(G,\vec{\theta})}}{Z(\vec{\theta})}, \tag{2}$$

being H the graph Hamiltonian such that $H(G,\vec{\theta}) = \vec{\theta} \cdot \vec{C(G)}$, and Z the normalizing partition function $Z = \sum e^{-H(G)}$. The set of variables $\vec{\theta}$ are the *Lagrange multipliers*, resulting from the maximization of Eq. 1 under the chosen constraints $\vec{C}$.

Secondly, we proceed to calculate the exact values of the Lagrange multipliers. Following Squartini and Garlaschelli [20] [22], we determine these parameters by imposing that the real network is found in the ensemble with maximum probability. Indeed, we may write the log-likelihood of observing the real network $L(\vec{\theta}) = ln(P(G^* \mid \vec{\theta}))$ as:

$$L(\vec{\theta}) = -H(G^*,\vec{\theta}) - lnZ(\vec{\theta}). \tag{3}$$

Maximizing this quantity thus allows fixing the $\vec{\theta}^*$ values. This second requirement ensures not only that the constraints are met on average, but also that they are the most likely ones, which is a warranty of non-bias [20].

Ensemble for a bipartite network with constrained degree sequences

We now explain how the randomizing scheme by Squartini and Garlaschelli [20] applies to our specific problem, namely a *bipartite network* subject to *local* constraints. The scheme has already been applied to study international trade networks [21].

To begin with, we construct the hamiltonian for a bipartite network, whose bipartite matrix we call $\mathbf{B}$. At variance with the monopartite case, we have *two* degree sequences (one for each of the guilds) which need to be taken into account separately. Although the scheme is equally valid for any mutualistic network (seed-dispersers, ant-plants...), for the sake of clarity we restrict our notation to the paradigmatic case of plant-pollinator communities. Thus, we will speak of systems with N_P number of *plants* and N_A pollinating *animals*. The constrained degree sequences, given by the real network, will be represented respectively by $\vec{v}$ and $\vec{h}$, where v_p is the diversity of *visiting* animal species that a plant species p receives, while h_a is the number of different *hosting* plant species with which a pollinator species a interacts.

In order to enforce both distributions as constraints we introduce two sets of Lagrange multipliers, $\vec{\alpha}$ for plants and $\vec{\beta}$ for animals. Subsequently, the graph hamiltonian can be written as

$$H(\mathbf{B},\vec{\alpha},\vec{\beta}) = \vec{\alpha} \cdot \vec{v} + \vec{\beta} \cdot \vec{h} \tag{4}$$

This means that the probability, Eq. 2, of encountering a bipartite graph B in the exponential random graph ensemble becomes:

$$P(\mathbf{B} \mid \vec{\alpha},\vec{\beta}) = \frac{e^{-\vec{\alpha}\cdot\vec{v}-\vec{\beta}\cdot\vec{h}}}{\sum_B e^{-\vec{\alpha}\cdot\vec{v}-\vec{\beta}\cdot\vec{h}}} \tag{5}$$

To simplify the notation we introduce the variable change $x_p = e^{-\alpha_p}$ and $y_a = e^{-\beta_a}$, as suggested as well by Squartini and Garlaschelli. Then, the log-likelihood of encountering the real network is:

$$L(\vec{x}) = \sum_{p=1}^{N_P} v_p \, ln(x_p) + \sum_{a=1}^{N_A} h_a \, ln(y_a) - \sum_{a=1}^{N_A} \sum_{p=1}^{N_P} ln(1 + x_p y_a) \tag{6}$$

which we need to maximize in order to find the optimal variables $\vec{x}^*$ and $\vec{y}^*$, that ultimately define our ensemble. Indeed, by requiring that $\vec{\nabla} L(\vec{x}, \vec{y}) = \vec{0}$, we obtain the following set of equations:

$$v_p = \sum_{a=1}^{N_A} \frac{x_p y_a}{1 + x_p y_a} \quad \text{for } p = 1, ..., N_P \tag{7}$$

$$h_a = \sum_{p=1}^{N_P} \frac{x_p y_a}{1 + x_p y_a} \quad \text{for } a = 1, ..., N_A \tag{8}$$

It can be easily shown that these equations are equivalent to imposing that the average degrees (right hand side) are equal to the degree sequence from the real network (left hand side), as we do below.

Probability matrix of interactions

Garlaschelli and Squartini also showed [20] that in the case of local constraints, the ensemble probability can be factorized, using our notation, in terms of the *probability of existence of a link between a plant species 'p' and animal species 'a'*, which we call p_{pa}. In effect, by taking $p_{pa} = \frac{x_p y_a}{1 + x_p y_a}$, replacing it into equation 5 and doing some little algebra, one finds:

$$P(\mathbf{B} \mid \vec{\alpha}, \vec{\beta}) = \prod_{p,a} {p_{pa}}^{b_{pa}} \, (1 - p_{pa})^{1 - b_{pa}} \tag{9}$$

Where b_{pa} is the (p,a) element of the bipartite matrix of interactions. Then, using expression 9, it is almost immediate to see that $\langle b_{pa} \rangle = p_{pa}$, thus in turn, $\langle b_{pa} \rangle = \frac{x_p y_a}{1 + x_p y_a}$. This shows that, as we had said, the right hand-side of equations 7-8 is a sum over a column or row of expected values of the randomized bipartite matrix.

We also note that the possibility of factorizing $P(\mathbf{B} \mid \vec{\alpha}, \vec{\beta})$ essentially entails that the probabilities p_{pa} are independent among them. In other words, when the constraints enforced are local, the probability of existence of different links are independent among them. This automatically allows the construction of the exact expected randomized matrix of interactions:

$$\langle \mathbf{B}^* \rangle = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1a} & \cdots & p_{1N_A} \\ p_{21} & p_{22} & \cdots & p_{1a} & \cdots & p_{2N_A} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{p1} & p_{p2} & \cdots & p_{pa} & \cdots & p_{pN_A} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{N_P 1} & p_{N_P 2} & \cdots & p_{N_P a} & \cdots & p_{N_P N_A} \end{pmatrix} \tag{10}$$

Computational implementation

Here we give the numerical details on how we obtained the Lagrange multipliers $\vec{x}^*$ and $\vec{y}^*$ that eventually define the corresponding statistical ensembles of the empirical networks. As proposed by Squartini and Garlaschelli [20], encountering these multipliers might be achieved following either of two procedures: by directly maximizing the log-likelihood in Eq. 6 through an optimizing algorithm, or by solving the non-linear, coupled set of equations in 7-8.

First, we optimized the log-likelihood by means of a global search, pseudo-random algorithm belonging to the Monte-Carlo family and known as *simulated annealing* [32–34]. In short, this method aims to find the global minimum of a function by sequentially exploring the solution space, through producing random proposals subjected to an accepting criteria. More specifically, it uses the Metropolis criteria [35], which is driven by a parameter T traditionally called *temperature*. The algorithm works iteratively, by finding the most probable state at each temperature and then using it as the initial condition in the next step. We start with a high temperature, which means that almost all proposals are accepted. As the algorithm progresses, the temperature decreases and so does the acceptance probability. This makes the search to become more and more restricted around eventual solutions, until a certain number of consecutive iterations (in our case, five) have produced solutions differing less than a certain tolerance, called tol. When that happens, we are confident enough of having reached the ground state and the algorithm stops. Given the pseudo-aleatory character of this approach, which allows escaping from local hills, it is extendedly used in situations in which the co-existence of several local optima is suspected.

In our case, since the algorithm is originally intended for minimizing but we aim to maximize, we simply took the minus of the function. Additionally, instead of optimizing the log-likelihood as written in Eq. 6, we incorporated the fact that the degrees may be degenerate. This means that nodes of the same guild having identical degrees satisfy equivalent equations, hence necessarily bearing the same solution. To account for this, we introduced a multiplicity factor m_p for plants and m_a for animals. If we call red_P and red_A the redundancy for plants and for animals (namely, the corresponding numbers of repeated degrees), then the system can be redimensionalized to $N'_P = N_P - red_P$ and $N'_A = N_A - red_A$. Consequently the log-likelihood might be rewritten into:

$$L(\vec{x}) = \sum_{p=1}^{N'_P} m_p v_p \, ln(x_p) + \sum_{a=1}^{N'_A} m_a h_a \, ln(y_a) - \sum_{a=1}^{N'_A} \sum_{p=1}^{N'_P} m_p m_a \, ln(1 + x_p y_a) \tag{11}$$

Although, in analytical terms, the original expression in Eq. 6 and this latter one are obviously equivalent, from a computational point of view reducing the number of variables enhances the algorithm's efficiency. Besides, imposing from the beginning such identity between variables improves the accuracy of the program.

We programmed a classical version of simulated annealing, with a starting temperature of $T = 10^3$, a reduction factor of the temperature of $RT = 0.85$, a tolerance $tol = 10^{-6}$ and a total number of updates per fixed temperature of $2 \cdot 10^4$. Furthermore, we ran the algorithm 10 times per network with different random seeds, in order to produce independent sequences of explorations. When all runs convergence to the same solution, as outlined by Goffe [34], it is extremely probable that we have certainly encountered the global optimum.

Secondly, we solved the set of equations by means of a local, deterministic algorithm known as the *modified Powell hybrid method*. In particular, we used the MINPACK library [36] for FORTRAN, available online [37]. This method finds the zero of a non-linear system by exploiting its Jacobian, which we analytically calculated and implemented into the program.

Like before, we re-dimensionalized the problem to N'_P equations for plants and N'_A equations for animals, which now read:

$$v_p = \sum_{a=1}^{N'_A} \frac{m_a x_p y_a}{1 + x_p y_a} \quad \text{for } p = 1, ..., N'_P \tag{12}$$

$$h_a = \sum_{p=1}^{N'_P} \frac{m_p x_p y_a}{1 + x_p y_a} \quad \text{for } a = 1, ..., N'_A \tag{13}$$

We implemented these equations and their Jacobian and ran the algorithm with a tolerance $tol = 10^{-11}$ (as defined in the source code). The possibility of exploiting the gradient provides, in general, a greater local accuracy than the simulated annealing technique. However, its shortcoming lays in the risk of getting trapped in local optima, from which, due to its deterministic nature, it is unable to escape. To compensate this drawback we performed a significant sampling of the space of initial conditions, by running 10^4 iterations of the algorithm, each with a different random selection of starting points, covering as well distinct ranges. However, due to the encounter of rough, rather accidental configuration surfaces, the modified Powell hybrid method was not always able to converge to a solution. The rate of success was approximately 50%·

To finally ensure that we found the global maximum, we compared the outcomes of the various independent runs and also, when the Powell algorithm succeeded, among both methods (so in total 10 runs for the simulated annealing and 10^4 for the Powell hybrid method). We confirmed, in all cases, that the same maxima was found. This enables us to assume that we found the actual *global* optimizing Lagrange Multipliers for each one of the networks of our study.

Moreover, the constraints were successfully met with a relative precision between 0.01% and 10%. This check was carried on by computing the average degrees using equations 7-8 and comparing the output with the imposed degree sequences (extracted from the empirical networks). The worst case of 10% was typically caused by discrepancies in low degrees, generally the most sensitive to imprecisions in the elements of the randomized matrix (since the matrix elements of low degree nodes are usually very small, see Fig.1 in main text as an example). Altogether, this second check warrants that our constrained optimization worked as aimed.

S2. Nestedness statistical measures using *NODF*

Here we describe how to obtain statistical measures in the random ensemble through analytical expressions, and particularly present our derivation for the nestedness metric known as NODF [38].

General analytical expressions

Let us call a property by X and its randomized measure (that is, the average across the random ensemble) by $\langle X \rangle^*$. When the property X can be calculated through an *analytical* expression (non-algorithmic) as a function of the bipartite matrix $\mathbf{B}$,

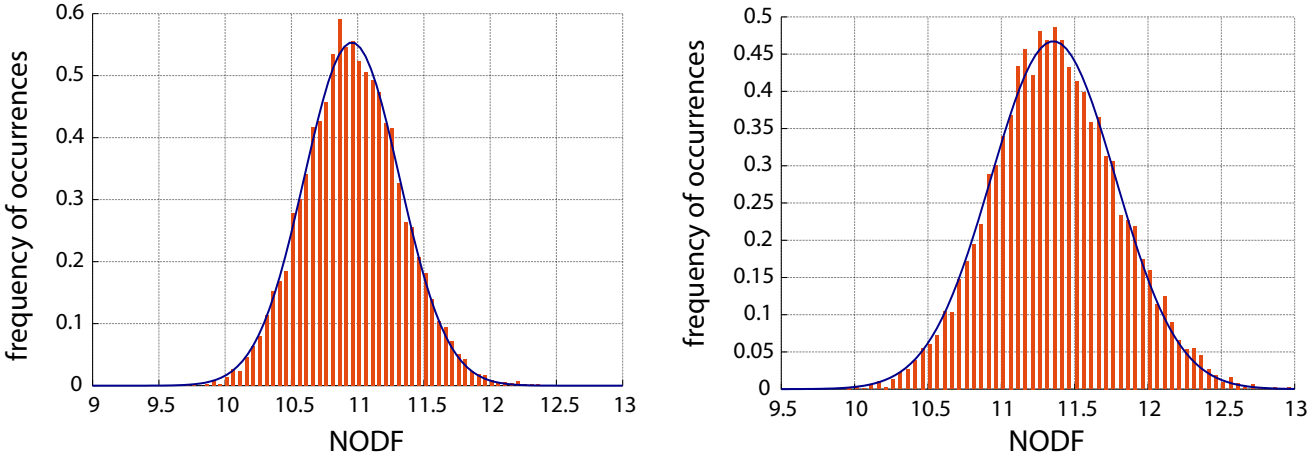

Figure 4. Nestedness distribution for two samplings of the statistical ensembles corresponding to the empirical networks by Petanidou et al. [39] (left) and by Inoue et al. [23] (right). In blue, fits of a gaussian function using the mean and standard deviation extracted from each distribution.

then Squartini and Garlaschelli [20] showed that it is possible to perform an approximate but accurate measure of the first and second moments of X, *directly* on $\langle \mathbf{B}^* \rangle$ (see Eq. 10, SI1). In particular, for the bipartite case, this reads:

$$\langle X \rangle^* \simeq X(\langle \mathbf{B}^* \rangle) \tag{14}$$

$$\sigma_X \simeq \sqrt{\sum_{p=1}^{N_P} \sum_{a=1}^{N_A} \left(\frac{\partial X(\mathbf{B})}{\partial b_{pa}} \right)^2 \sigma_{b_{pa}}^2} \tag{15}$$

Where $\sigma_{b_{pa}}$ is the standard deviation for the bipartite matrix element b_{pa}. The condition for these approximations to be accurate is that the property X be gaussian-distributed in the random ensemble.

Derivation for *NODF*

Let us see how the previous expressions can be applied to measure nestedness with the *nestedness metric based on overlap and decreasing fill* (hereafter, *NODF*) by Almeida-Neto et al. [38].

We first verified that the assumption of gaussianity is fulfilled by performing a check on a smaller subset of our set of empirical networks. To do this, for each of the corresponding statistical ensembles we generated a sample of 10^4 networks obeying the probability of link existence given by $\langle \mathbf{B}^* \rangle$ (see Eq. 10, SI1). We then computed the nestedness of each sampled network in order to generate the nestedness distribution. In all cases we could successfully fit a gaussian function (see Fig. 5 as an example).

Once shown that the gaussian behavior is satisfied for this metric, we next apply Eqs. 14-15. Yet, we need an analytical, packed expression that facilitates the calculations of the metric. The NODF basically considers two contributing factors to nestedness: *decreasing fill* (the fact that, after a proper ordering, both degree sequences progressively decline) and *paired overlap* (the number of shared partners between two columns or rows, normalized by the smaller degree). By gathering together the sequential analysis indicated by Almeida-Neto et al., we proposed a novel compact expression to calculate NODF, that reads:

$$\mathrm{NODF}(\mathbf{B}) = \frac{1}{K} \sum_{i<j}^{N_P} \left\{ [1-\theta(v_j - v_i)] \cdot \frac{\sum_{a=1}^{N_A} b_{ia} b_{ja}}{v_j} \right\} + \frac{1}{K} \sum_{k<l}^{N_A} \left\{ [1-\theta(h_l - h_k)] \cdot \frac{\sum_{p=1}^{N_P} b_{pk} b_{pl}}{h_l} \right\} \tag{16}$$

$$\text{where } K^{-1} = \frac{N_P(N_P-1) + N_A(N_A+1)}{200} \tag{17}$$

We maintain here our previous notation (see SI1), so v_p is the degree of plant p and h_a the degree of animal a. The double sums run over two indices such that, as seen in Fig. 5, row i is placed upper row j and column k more to the left than column l.

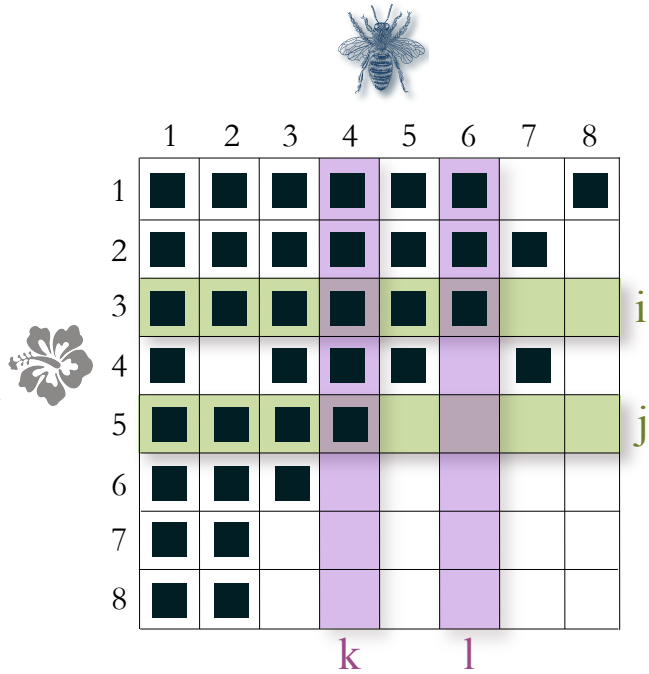

Figure 5. Example of an ordered matrix of interactions, not perfectly nested. Species of both guilds have been ordered in decreasing degree, and the numbered labels indicate their rank (the larger the degree, the smaller the rank). The indexes i, j, k and l illustrate our notation for rows and columns.

The K factor contains the normalization over all possible pairs, as well as a percentage rescaling (note that NODF takes values between 0 and 100). Finally, the θ term is the Heaviside step function, which is zero when its argument is negative, and one if its argument is positive or zero. In our context it serves to encapsulate the decreasing fill condition. In fact, from now on we will use the following abbreviations:

$$DF_{ij} = 1 - \theta(v_j - v_i) \quad \text{such that, if } v_j \geq v_i \text{ then } DF_{ij} = 0, \quad \text{and if } v_j < v_i \text{ then } DF_{ij} = 1 \tag{18}$$

$$DF_{kl} = 1 - \theta(h_l - h_k) \quad \text{such that, if } h_l \geq h_k \text{ then } DF_{kl} = 0, \quad \text{and if } h_l < h_k \text{ then } DF_{kl} = 1 \tag{19}$$

The analytical and packed expression for NODF that appears in Eq. 16 can then be plugged into Eqs. 14-15. Accordingly, we obtained that the first moment of the randomized NODF for a certain real bipartite matrix $\mathbf{B}^*$ reads:

$$\langle \text{NODF}(\mathbf{B}) \rangle^* = \frac{1}{K} \sum_{i<j}^{N_P} \left\{ DF_{ij} \cdot \frac{\sum_{a=1}^{N_A} \langle b_{ia} \rangle \langle b_{ja} \rangle}{\sum_{a=1}^{N_A} \langle b_{ja} \rangle} \right\} + \frac{1}{K} \sum_{k<l}^{N_A} \left\{ DF_{kl} \cdot \frac{\sum_{p=1}^{N_P} \langle b_{pk} \rangle \langle b_{pl} \rangle}{\sum_{p=1}^{N_P} \langle b_{pl} \rangle} \right\} \tag{20}$$

Notice that $\sum_{a=1}^{N_A} \langle b_{pa} \rangle = v_p$ and $\sum_{p=1}^{N_P} \langle b_{pa} \rangle = h_a$, given that the randomized matrix necessarily fulfills the enforced constrains. Additionally, this warrants that the ordering of the matrix is equal to the original one, which is important since NODF is ordering-dependent through the decreasing fill terms.

It is also interesting to remark that the previous expression can be understood in probabilistic terms. Indeed, given that $\langle b_{pa} \rangle = p_{ap}$, where p_{pa} are independent link probabilities, the overlap term might be seen as a joint probability of two independent events, divided by a normalizing factor which is the union of independent probabilities. For example, for one pair of animals, the overlap term results in:

$$\frac{\sum_{p=1}^{N_P} \langle b_{pk} \rangle \langle b_{pl} \rangle}{\sum_{p=1}^{N_P} \langle b_{pl} \rangle} = \frac{\sum_{p=1}^{N_P} p_{pk}\, p_{pl}}{\sum_{p=1}^{N_P} p_{pl}} = \frac{\sum_{p=1}^{N_P} P(p_{pk} \cap p_{pl})}{P(p_{1l} \cup p_{2l} \cup \ldots\, p_{pl} \cup \ldots\, p_{N_P l})} \tag{21}$$

fraction of ntws with |z-score| ≤ 1		fraction of ntws with |z-score| ≤ 2	
84 out of 167	50.3%	149 out of 167	89.2%

Table 2. Fraction of networks whose discrepancy between the real and randomized nestedness is less or equal than one or two sigma.

fraction of ntws with corrected |z-score| ≤ 1		fraction of ntws with corrected |z-score| ≤ 2	
151 out of 167	90.4%	161 out of 167	96.4%

Table 3. After performing the multiple test correction using the false discovery rate method (see Methods summary in Main text), fraction of networks whose discrepancy between the real and randomized nestedness is less or equal than one or two sigma.

Apart from the first moment, equation 15 indicates how to compute the standard deviation. Adapting it to the nestedness measure, we encounter:

$$\sigma_{\text{NODF}} = \sqrt{\sum_{p=1}^{N_P}\sum_{a=1}^{N_A}\left(\frac{\partial \text{NODF}(\mathbf{B})}{\partial b_{pa}}\right)^2 \sigma^2_{b_{pa}}} \qquad \text{where} \quad \sigma^2_{b_{pa}} = p_{pa}\,(1-p_{pa}) \tag{22}$$

We have introduced that the standard deviation of the probability of links is that of a Bernoulli distribution, since the existence of a link is a binomial process (either there is interaction or there is not). Furthermore, the derivative with respect to a general matrix element b_{rc} (the sub-index r stands for rows and c stands for columns) can be split into the contributions of plants and of animals:

$$\frac{\partial\ \text{NODF}(\mathbf{B})}{\partial b_{rc}} = \frac{\partial\ \text{NODF}(\mathbf{B})_{\text{plants}}}{\partial b_{rc}} + \frac{\partial\ \text{NODF}(\mathbf{B})_{\text{animals}}}{\partial b_{rc}} \tag{23}$$

After deriving, we obtained that:

$$\frac{\partial\ \text{NODF}(\mathbf{B})_{\text{plants}}}{\partial b_{rc}} = \frac{1}{K}\left\{\sum_{j=r+1}^{N_P} DF_{rj}\,\frac{b_{jc}}{v_j} + \sum_{i=1}^{r-1} DF_{ir}\,\frac{b_{ic}}{v_r} - \sum_{i=1}^{r-1}\sum_{a=1}^{N_A} DF_{ir}\,\frac{b_{ia}\,b_{ra}}{{v_r}^2}\right\} \tag{24}$$

$$\frac{\partial\ \text{NODF}(\mathbf{B})_{\text{animals}}}{\partial b_{rc}} = \frac{1}{K}\left\{\sum_{l=c+1}^{N_A} DF_{cl}\,\frac{b_{rl}}{h_l} + \sum_{k=1}^{c-1} DF_{kc}\,\frac{b_{rk}}{h_c} - \sum_{k=1}^{c-1}\sum_{p=1}^{N_P} DF_{kc}\,\frac{b_{pk}\,b_{pc}}{{h_c}^2}\right\} \tag{25}$$

These novel estimations of the first and second moments can then be used to provide a randomized measure of nestedness $\langle \text{NODF}(\mathbf{B})\rangle^*$ together with its statistical significance.

S3. Nestedness statistical measures using the *largest eigenvalue radius*

The largest eigenvalue radius was recently proposed by Staniczenko et al. [12] as an alternative measure for nestedness that directly relies on the spectral properties of the adjacency matrix. The fact that it involves finding the *maximum* eigenvalue entails that we lack an analytical and derivable expression for it. This, in turn, means that we can not derive for this metric the analytical expressions of Eqs. 14-15 in SI2. Therefore, in order to calculate the statistical properties of our interest, we produced a sample of the statistical ensemble and algorithmically computed the distribution of the largest eigenvalue radius.

In detail, we produced 10^4 networks, sampled using the link probabilities in $\langle \mathbf{B}^* \rangle$ (see Eq. 10, SI1). Then we computed the largest eigenvalue radius, which we call $\rho(\lambda)$, using the *R* package *rARPACK* [40]. Finally we calculated the average and the standard deviation of the resulting distribution. As shown in Fig. 6, the results are fully compatible with those found using the NODF nestedness metric. See also Tables. 2 and 3 for details of percentages.

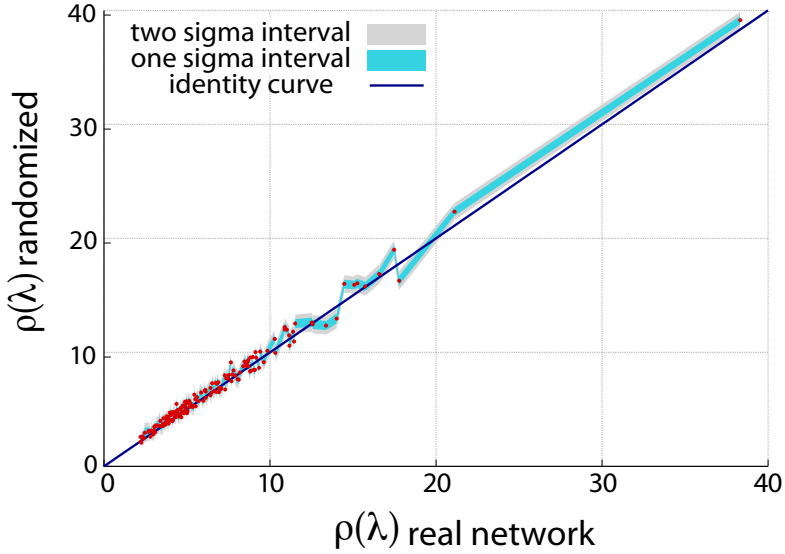

Figure 6. Comparison between the real observation of the largest eigenvalue radius $\rho(\lambda)$ and the average estimation in the statistical ensemble, for the 167 networks of our study.

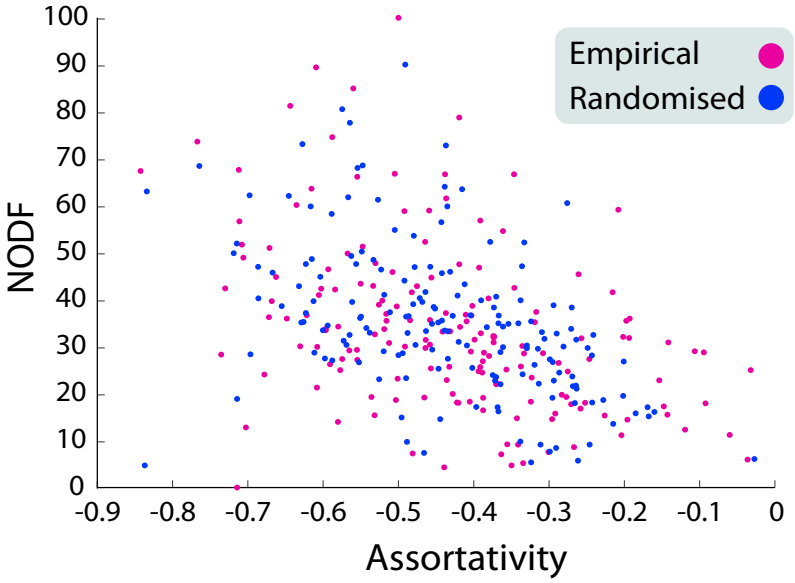

Figure 7. Relation between the degree assortativity and the nestedness for the real network (in red) and the average measure for the randomised case (in blue).

S4. Assortativity statistical measures

Assortativity is a network feature that quantifies to what extent nodes tend to match other nodes that are similar (or dissimilar) to them. Here, we particularly used the notion of degree assortativity, which means that *similarity* is labelled by the degree. We followed the definition proposed by Newman [41], which consists of a normalized correlation coefficient between degrees. This eventually corresponds to the *Pearson correlation coefficient* denoted by r, such that $r = -1$ indicates perfect disassortativity, $r = 0$ no correlation at all and $r = 1$ maximum assortativity.

In order to compute the statistical properties of this quantity, we produced for each ensemble a sampling made up by 10^4 networks. We then measured computationally the assortativity of each sampled network using the *assortativity_degree* function from the *igraph* package in *R* [42]. Finally this allowed us to calculate the first and second moments of the assortativity for each ensemble in our set.

S5. Data

In our study we analyzed 167 real interaction networks from the *Web of Life* dataset [43]. This set consists of 133 plant-pollinator communities [references from 44 to 80], 30 seed-dispersal [references from 81 to 101] and 4 plant-ant [references from 102 to 105].

Data sometimes included information about link's weight, but we converted all networks to binary matrices.

References

[1] J. Bascompte, P. Jordano, C. J. Melián, and J. M. Olesen, "The nested assembly of plant–animal mutualistic networks," *Proceedings of the National Academy of Sciences*, vol. 100, no. 16, pp. 9383–9387, 2003.

[2] U. Bastolla, M. A. Fortuna, A. Pascual-García, A. Ferrera, B. Luque, and J. Bascompte, "The architecture of mutualistic networks minimizes competition and increases biodiversity," *Nature*, vol. 458, no. 7241, pp. 1018–1020, 2009.

[3] E. Thébault and C. Fontaine, "Stability of ecological communities and the architecture of mutualistic and trophic networks," *Science*, vol. 329, no. 5993, pp. 853–856, 2010.

[4] R. P. Rohr, S. Saavedra, and J. Bascompte, "On the structural stability of mutualistic systems," *Science*, vol. 345, no. 6195, p. 1253497, 2014.

[5] E. Burgos, H. Ceva, R. P. Perazzo, M. Devoto, D. Medan, M. Zimmermann, and A. M. Delbue, "Why nestedness in mutualistic networks?," *Journal of theoretical biology*, vol. 249, no. 2, pp. 307–313, 2007.

[6] S. Suweis, F. Simini, J. R. Banavar, and A. Maritan, "Emergence of structural and dynamical properties of ecological mutualistic networks," *Nature*, vol. 500, no. 7463, pp. 449–452, 2013.

[7] S. Valverde, J. Montoya, L. Joppa, and R. Sole, "Is nestedness in mutualistic networks an evolutionary spandrel?," *ArXiv preprint*.

[8] W. Ulrich, M. Almeida-Neto, and N. J. Gotelli, "A consumer's guide to nestedness analysis," *Oikos*, vol. 118, no. 1, pp. 3–17, 2009.

[9] D. P. Vázquez and M. A. Aizen, "Community-wide patterns of specialization in plant–pollinator interactions revealed by null models," *Plant–pollinator interactions: from specialization to generalization*, pp. 200–219, 2006.

[10] E. L. Rezende, P. Jordano, and J. Bascompte, "Effects of phenotypic complementarity and phylogeny on the nested structure of mutualistic networks," *Oikos*, vol. 116, no. 11, pp. 1919–1929, 2007.

[11] R. P. Perazzo, L. Hernández, H. Ceva, E. Burgos, and J. I. Alvarez-Hamelin, "Study of the influence of the phylogenetic distance on the interaction network of mutualistic ecosystems," *Physica A: Statistical Mechanics and its Applications*, vol. 394, pp. 124–135, 2014.

[12] P. P. Staniczenko, J. C. Kopp, and S. Allesina, "The ghost of nestedness in ecological networks," *Nature communications*, vol. 4, p. 1391, 2013.

[13] A. James, J. W. Pitchford, and M. J. Plank, "Disentangling nestedness from models of ecological complexity," *Nature*, vol. 487, no. 7406, pp. 227–230, 2012.

[14] S. Jonhson, V. Domínguez-García, and M. A. Muñoz, "Factors determining nestedness in complex networks," *PloS one*, vol. 8, no. 9, p. e74025, 2013.

[15] W. Feng and K. Takemoto, "Heterogeneity in ecological mutualistic networks dominantly determines community stability," *Scientific reports*, vol. 4, p. 5912, 2014.

[16] D. Medan, R. P. Perazzo, M. Devoto, E. Burgos, M. G. Zimmermann, H. Ceva, and A. M. Delbue, "Analysis and assembling of network structure in mutualistic systems," *Journal of theoretical biology*, vol. 246, no. 3, pp. 510–521, 2007.

[17] W. Atmar and B. D. Patterson, "The measure of order and disorder in the distribution of species in fragmented habitat," *Oecologia*, vol. 96, no. 3, pp. 373–382, 1993.

[18] J. M. Olesen, J. Bascompte, Y. L. Dupont, H. Elberling, C. Rasmussen, and P. Jordano, "Missing and forbidden links in mutualistic networks," *Proceedings of the Royal Society of London B: Biological Sciences*, pp. 13–71, 2010.

[19] L. N. Joppa, J. M. Montoya, R. Solé, J. Sanderson, and S. L. Pimm, "On nestedness in ecological networks," *Evolutionary Ecology Research*, vol. 12, no. 1, pp. 35–46, 2010.

[20] T. Squartini and D. Garlaschelli, "Analytical maximum-likelihood method to detect patterns in real networks," *New Journal of Physics*, vol. 13, no. 8, p. 083001, 2011.

[21] F. Saracco, R. Di Clemente, A. Gabrielli, and T. Squartini, "Randomizing bipartite networks: the case of the world trade web," *Scientific Reports*, vol. 5, no. 10595, 2015.

[22] D. Garlaschelli and M. I. Loffredo, "Maximum likelihood: extracting unbiased information from complex networks," *Physical Review E*, vol. 78, no. 1, p. 015101, 2008.

[23] T. Inoue, M. Kato, T. Kakutani, T. Suka, and T. Itino, "Insect-flower relationship in the temperate deciduous forest of kibune, kyoto: an overview of the flowering phenology and the seasonal pattern of insect visits," 1990.

[24] W. Ulrich and N. J. Gotelli, "Null model analysis of species nestedness patterns," *Ecology*, vol. 88, no. 7, pp. 1824–1831, 2007.

[25] P. Jordano, J. Bascompte, and J. M. Olesen, "Invariant properties in coevolutionary networks of plant–animal interactions," *Ecology Letters*, vol. 6, no. 1, pp. 69–81, 2003.

[26] J. Memmott, N. M. Waser, and M. V. Price, "Tolerance of pollination networks to species extinctions," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 271, no. 1557, pp. 2605–2611, 2004.

[27] N. J. Gotelli and G. L. Entsminger, "Swap and fill algorithms in null model analysis: rethinking the knight's tour," *Oecologia*, vol. 129, no. 2, pp. 281–291, 2001.

[28] S. Johnson, J. J. Torres, J. Marro, and M. A. Muñoz, "Entropic origin of disassortativity in complex networks," *Physical review letters*, vol. 104, no. 10, p. 108702, 2010.

[29] K. Takemoto and M. Arita, "Nested structure acquired through simple evolutionary process," *Journal of theoretical biology*, vol. 264, no. 3, pp. 782–786, 2010.

[30] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the royal statistical society. Series B (Methodological)*, pp. 289–300, 1995.

[31] J. Park and M. E. Newman, "Statistical mechanics of networks," *Physical Review E*, vol. 70, no. 6, p. 066117, 2004.

[32] A. Corana, M. Marchesi, C. Martini, and S. Ridella, "Minimizing multimodal functions of continuous variables with the "simulated annealing" algorithm corrigenda for this article is available here," *ACM Transactions on Mathematical Software (TOMS)*, vol. 13, no. 3, pp. 262–280, 1987.

[33] W. L. Goffe, G. D. Ferrier, and J. Rogers, "Global optimization of statistical functions with simulated annealing," *Journal of econometrics*, vol. 60, no. 1-2, pp. 65–99, 1994.

[34] W. L. Goffe *et al.*, "Simann: A global optimization algorithm using simulated annealing," *Studies in Nonlinear Dynamics and Econometrics*, vol. 1, no. 3, pp. 169–176, 1996.

[35] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *The journal of chemical physics*, vol. 21, no. 6, pp. 1087–1092, 1953.

[36] J. J. Moré, B. S. Garbow, and K. E. Hillstrom, "User guide for minpack-1," tech. rep., CM-P00068642, 1980.

[37] J. J. M. Burton S. Garbow, Kenneth E. Hillstrom, *MINPACK libary source code, subroutine HYBRD*. Available at: `https://www.math.utah.edu/software/minpack/minpack/hybrd.html`.

[38] M. Almeida-Neto, P. Guimarães, P. R. Guimarães, R. D. Loyola, and W. Ulrich, "A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement," *Oikos*, vol. 117, no. 8, pp. 1227–1239, 2008.

[39] T. Petanidou, "Pollination ecology in a phryganic ecosystem," 1991.

[40] *rARPACK: Solvers for Large Scale Eigenvalue and SVD Problems*. Available at: `https://CRAN.R-project.org/package=rARPACK`.

[41] M. E. Newman, "Assortative mixing in networks," *Physical review letters*, vol. 89, no. 20, p. 208701, 2002.

[42] G. Csardi, *Assortativity in R igraph package*. Available at: `http://igraph.org/r/doc/assortativity.html`.

[43] Bascompte Lab, *Web of Life, ecological networks database*. Available at `http://www.web-of-life.es/`.

[44] M. T. K. Arroyo, R. Primack, and J. Armesto, "Community studies in pollination ecology in the high temperate andes of central chile. i. pollination mechanisms and altitudinal variation," *American journal of botany*, pp. 82–97, 1982.

[45] S. C. Barrett and K. Helenurm, "The reproductive biology of boreal forest herbs. i. breeding systems and pollination," *Canadian Journal of Botany*, vol. 65, no. 10, pp. 2036–2046, 1987.

[46] F. E. Clements and F. L. Long, *Experimental pollination: an outline of the ecology of flowers and insects*. No. 336, Carnegie institution of Washington, 1923.

[47] L. Dicks, S. Corbet, and R. Pywell, "Compartmentalization in plant–insect flower visitor webs," *Journal of Animal Ecology*, vol. 71, no. 1, pp. 32–43, 2002.

[48] Y. L. Dupont, D. M. Hansen, and J. M. Olesen, "Structure of a plant–flower-visitor network in the high-altitude sub-alpine desert of tenerife, canary islands," *Ecography*, vol. 26, no. 3, pp. 301–310, 2003.

[49] H. Elberling and J. M. Olesen, "The structure of a high latitude plant-flower visitor system: the dominance of flies," *Ecography*, vol. 22, no. 3, pp. 314–323, 1999.

[50] J. M. Olesen, L. I. Eskildsen, and S. Venkatasamy, "Invasion of pollination networks on oceanic islands: importance of invader complexes and endemic super generalists," *Diversity and Distributions*, vol. 8, no. 3, pp. 181–192, 2002.

[51] J. Ollerton, S. D. Johnson, L. Cranmer, and S. Kellie, "The pollination ecology of an assemblage of grassland asclepiads in south africa," *Annals of Botany*, vol. 92, no. 6, pp. 807–834, 2003.

[52] B. Hocking, "Insect-flower associations in the high arctic with special reference to nectar," *Oikos*, pp. 359–387, 1968.

[53] J. Herrera, "Pollination relationships in southern spanish mediterranean shrublands," *The Journal of Ecology*, pp. 274–287, 1988.

[54] J. Memmott, "The structure of a plant-pollinator food web," *Ecology letters*, vol. 2, no. 5, pp. 276–280, 1999.

[55] D. W. Inouye and G. H. Pyke, "Pollination biology in the snowy mountains of australia: comparisons with montane colorado, usa," *Austral Ecology*, vol. 13, no. 2, pp. 191–205, 1988.

[56] P. G. Kevan, "High arctic insect-flower relations the interrelationships of arthropods and flowers at lake hazen, ellesmere island, nwt, canada.," 2002.

[57] M. Kato, T. Kakutani, T. Inoue, and T. Itino, "Insect-flower relationship in the primary beech forest of ashu, kyoto: an overview of the flowering phenology and the seasonal pattern of insect visits," 1990.

[58] D. Medan, N. H. Montaldo, M. Devoto, A. Mantese, V. Vasellati, G. G. Roitman, and N. H. Bartoloni, "Plant-pollinator relationships at two altitudes in the andes of mendoza, argentina," *Arctic, Antarctic, and Alpine Research*, pp. 233–241, 2002.

[59] T. Mosquin and J. Martin, "Observations on the pollination biology of plants on melville island, nwt, canada," *Canadian Field Naturalist*, vol. 81, pp. 201–205, 1967.

[60] "Pollination ecology of the spring wildflower community in the deciduous forests of piedmont north carolina,"

[61] C. McMullen, "Flower-visiting insects of the galapagos islands," *Pan-Pacific Entomologist*, vol. 69, no. 1, pp. 95–106, 1993.

[62] R. B. Primack, "Insect pollination in the new zealand mountain flora," *New Zealand Journal of Botany*, vol. 21, no. 3, pp. 317–333, 1983.

[63] N. Ramirez and Y. Brito, "Pollination biology in a palm swamp community in the venezuelan central plains," *Botanical Journal of the Linnean Society*, vol. 110, no. 4, pp. 277–302, 1992.

[64] N. Ramirez, "Biologia de polinizacion en una comunidad arbustiva tropical de la alta guayana venezolana," *Biotropica*, pp. 319–330, 1989.

[65] D. W. Schemske, M. F. Willson, M. N. Melampy, L. J. Miller, L. Verner, K. M. Schemske, and L. B. Best, "Flowering ecology of some spring woodland herbs," *Ecology*, vol. 59, no. 2, pp. 351–366, 1978.

[66] E. Small, "Insect pollinators of the mer bleue peat bog of ottawa," *Canadian field-naturalist*, 1976.

[67] C. Smith-Ramirez, P. Martinez, M. Nunez, C. González, and J. J. Armesto, "Diversity, flower visitation frequency and generalism of pollinators in temperate rain forests of chiloé island, chile," *Botanical Journal of the Linnean Society*, vol. 147, no. 4, pp. 399–416, 2005.

[68] M. Percival, "Floral ecology of coastal scrub in southeast jamaica," *Biotropica*, pp. 104–129, 1974.

[69] A. Montero, *The Ecology of Three Pollination Networks*. PhD thesis, MSc thesis, Aarhus University, Aarhus, 2005.

[70] T. Ingversen, *Plant–pollinator interactions on Jamaica and Dominica: The centrality, asymmetry and modularity of networks*. PhD thesis, MSc thesis, Aarhus University, Aarhus, 2006.

[71] M. Philipp, J. Böcher, H. R Siegismund, and L. R Nielsen, "Structure of a plant-pollinator network on a pahoehoe lava desert of the galápagos islands," *Ecography*, vol. 29, no. 4, pp. 531–540, 2006.

[72] M. Kato, "Anthophilous insect community and plant-pollinator interactions on amami islands in the ryukyu archipelago, japan (original paper)," 2000.

[73] R. Lundgren and J. M. Olesen, "The dense and highly connected world of greenland's plants and their pollinators," *Arctic, Antarctic, and Alpine Research*, vol. 37, no. 4, pp. 514–520, 2005.

[74] M. Bundgaard, *Tidslig og rumlig variation i et plante-bestøvernetværk*. PhD thesis, MSc. thesis, Aarhus University, Aarhus., 2003.

[75] Y. L. Dupont and J. M. Olesen, "Ecological modules and roles of species in heathland plant–insect flower visitor networks," *Journal of Animal Ecology*, vol. 78, no. 2, pp. 346–353, 2009.

[76] S. Bek, *A pollination network from a Danish forest meadow*. PhD thesis, MSc. thesis, Aarhus University, Aarhus., 2006.

[77] L. Stald, A. Valido, and J. Olesen, *Struktur og dynamik i rum og tid at et bestøvningsnetværk på Tenerife, De Kanariske Øer*. PhD thesis, MSc-thesis, University of Aarhus, Denmark, 2003.

[78] D. P. Vázquez, "Interactions among introduced ungulates, plants, and pollinators: a field study in the temperate forest of the southern andes," 2002.

[79] W. P. PhD thesis, BSc thesis, University of Aarhus, Denmark, 1998.

[80] K. Yamazaki and M. Kato, "Flowering phenology and anthophilous insect community in a grassland ecosystem at mt. yufu, western japan," 2003.

[81] J. W. Baird, "The selection and use of fruit by birds in an eastern forest," *The Wilson Bulletin*, pp. 63–73, 1980.

[82] B. Beehler, "Frugivory and polygamy in birds of paradise," *The Auk*, pp. 1–12, 1983.

[83] T. A. Carlo, J. A. Collazo, and M. J. Groom, "Avian fruit preferences across a puerto rican forested landscape: pattern consistency and implications for seed removal," *Oecologia*, vol. 134, no. 1, pp. 119–131, 2003.

[84] F. Crome, "The ecology of fruit pigeons in tropical northern queensland.," *Wildlife Research*, vol. 2, no. 2, pp. 155–185, 1975.

[85] P. Frost, "Fruit-frugivore interactions in a south african coastal dune forest," 1980.

[86] M. Galetti and M. A. Pizo, "Fruit eating by birds in a forest fragment in southeastern brazil.," *Revista Brasileira de Ornitologia-Brazilian Journal of Ornithology*, vol. 4, no. 5, p. 9, 2013.

[87] B. K. Snow and D. Snow, "The feeding ecology of tanagers and honeycreepers in trinidad," *The Auk*, vol. 88, no. 2, pp. 291–322, 1971.

[88] B. K. Snow and D. Snow, "Birds and berries," 1988.

[89] A. Hamann and E. Curio, "Interactions among frugivores and fleshy fruit trees in a philippine submontane rainforest," *Conservation Biology*, vol. 13, no. 4, pp. 766–773, 1999.

[90] P. Jordano, "El ciclo anual de los paseriformes frugívoros en el matorral mediterráneo del sur de españa: importancia de su invernada y variaciones interanuales," *Ardeola*, vol. 32, no. 1, pp. 69–94, 1985.

[91] G. E. Kantak, "Observations on some fruit-eating birds in mexico," *The Auk*, vol. 96, no. 1, pp. 183–186, 1979.

[92] F. Lambert, "Fig-eating by birds in a malaysian lowland rain forest," *Journal of Tropical Ecology*, vol. 5, no. 4, pp. 401–412, 1989.

[93] C. E. Tutin, R. M. Ham, L. J. White, and M. J. Harrison, "The primate community of the lopé reserve, gabon: diets, responses to fruit scarcity, and effects on biomass," *American Journal of Primatology*, vol. 42, no. 1, pp. 1–24, 1997.

[94] A. L. Mack and D. D. Wright, "Notes on occurrence and feeding of birds at crater mountain biological research station, papua new guinea," *Emu*, vol. 96, no. 2, pp. 89–101, 1996.

[95] N. T. Wheelwright, W. A. Haber, K. G. Murray, and C. Guindon, "Tropical fruit-eating birds and their food plants: a survey of a costa rican lower montane forest," *Biotropica*, pp. 173–192, 1984.

[96] W. Silva, "Patterns of fruit-frugivore interactions in two atlantic forest bird communities of south-eastern brazil: implications for conservation," *Seed dispersal and frugivory: Ecology, evolution and conservation*, pp. 423–435, 2002.

[97] N. Noma, "Annual fluctuations of sapfruits production and synchronization within and inter species in a warm temperate forest on yakushima island," *Tropics*, vol. 6, pp. 441–449, 1997.

[98] J. Guitián, "Relaciones entre los frutos y los passeriformes en un bosque montano de la cordillera cantábrica occidental," *Univ. Santiago, Spain*, 1983.

[99] R. H. Heleno, J. A. Ramos, and J. Memmott, "Integration of exotic seeds into an azorean seed dispersal network," *Biological Invasions*, vol. 15, no. 5, pp. 1143–1154, 2013.

[100] B. Poulin, S. J. Wright, G. Lefebvre, and O. Calderón, "Interspecific synchrony and asynchrony in the fruiting phenologies of congeneric bird-dispersed plants in panama," *Journal of Tropical Ecology*, vol. 15, no. 2, pp. 213–227, 1999.

[101] M. Schleuning, N. Blüthgen, M. Flörchinger, J. Braun, H. M. Schaefer, and K. Böhning-Gaese, "Specialization and interaction strength in a tropical plant–frugivore network differ among forest strata," *Ecology*, vol. 92, no. 1, pp. 26–36, 2011.

[102] N. Blüthgen, N. E Stork, and K. Fiedler, "Bottom-up control and co-occurrence in complex communities: Honeydew and nectar determine a rainforest ant mosaic," *Oikos*, vol. 106, no. 2, pp. 344–358, 2004.

[103] D. W. Davidson, R. R. Snelling, and J. T. Longino, "Competition among ants for myrmecophytes and the significance of plant trichomes," *Biotropica*, pp. 64–73, 1989.

[104] D. Davidson and B. Fisher, "Symbiosis of ants with cecropia as a function of light regime," *Huxley, C, R,, Cutler, D, F ed (s). Ant-plant interactions. Oxford Univ. Press: Oxford, etc*, pp. 289–309, 1991.

[105] C. R. Fonseca and G. Ganade, "Asymmetries, compartments and null interactions in an amazonian ant-plant community," *Journal of Animal Ecology*, pp. 339–347, 1996.