Breaking the Spell of Nestedness: The Entropic Origin of Nestedness in Mutualistic Systems

Clàudia Payrató-Borràs,[1,2] Laura Hernández,[1] and Yamir Moreno[2,3,4,*]

[1]*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, Zaragoza 50018, Spain
[3]*Department of Theoretical Physics, Faculty of Sciences, University of Zaragoza, Zaragoza 50009, Spain*
[4]*ISI Foundation, Turin 10126, Italy*

Mutualistic interactions, those that are naturally beneficial for both interacting species, are recurrently found in ecosystems. Observations of natural systems show that if we draw mutualistic relationships as links between species, the resulting mutualistic network of interactions displays a widespread particular ordering called nestedness. In such an ordering, the mutualistic partners of a given species conform a subset of the partners of all species with larger degree, that is, of those species having more interactions. On the other hand, theoretical works show that a nested structure has a positive impact on a number of relevant features of mutualistic communities ranging from species coexistence to structural stability and biodiversity. However, how nestedness emerges and what are its determinants, are still open challenges that have led to multiple debates to date. Here we show, by applying a theoretical approach to the analysis of 167 real mutualistic networks, that nestedness is not an irreducibly macroscopic feature but an entropic consequence of the degree sequences (number of mutualistic interactions of each species). Remarkably, we find that an outstanding majority of the analyzed networks does not show statistically 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 node-dependent properties rather than on global quantities.

Subject Areas: Complex Systems,
Interdisciplinary Physics,
Statistical Physics

I. INTRODUCTION

Almost two decades ago, a seminal work by Bascompte *et al.* [1] revealed that nested patterns are ubiquitous in ecological systems. These patterns were disclosed thanks to the translation of ecological communities into ecological networks. On a network representation, nodes depict species and links capture observed ecological interactions, like mutualism or predation. Moreover, the network is said to be bipartite if it can be partitioned into two distinct sets in such a way that nodes within the same set do not interact among themselves. For such bipartite networks, a *perfectly nested* structure is defined by the fact that the interactions of a given node are invariably a subset of the interactions of all nodes with larger degree (see Fig. 1). That is, if **B** is the

*[*]yamir.moreno@gmail.com

biadjacency matrix of a bipartite network, the network will be perfectly nested only if both of the following conditions are true:

if $\mathbf{B}_{i,l} = 1$ then $\mathbf{B}_{i,k} = 1 \Leftrightarrow$ given a pair of columns k and l such that

$$\sum_{i}^{N_R} \mathbf{B}_{i,l} \leq \sum_{i}^{N_R} \mathbf{B}_{i,k}, \quad \forall \ i, \tag{1}$$

if $\mathbf{B}_{j,k} = 1$ then $\mathbf{B}_{i,k} = 1 \Leftrightarrow$ given a pair of rows i and j such that

$$\sum_{k}^{N_C} \mathbf{B}_{j,k} \leq \sum_{k}^{N_C} \mathbf{B}_{i,k}, \quad \forall \ k, \tag{2}$$

where N_R and N_C are, respectively, the number of rows and columns of the biadjacency matrix **B**. The conditions above translate into the fact that specialist species, that is, species with few interactions and thus a small degree, are seldom interacting with other specialists. Instead, they appear attached to generalist species, which have a large degree,
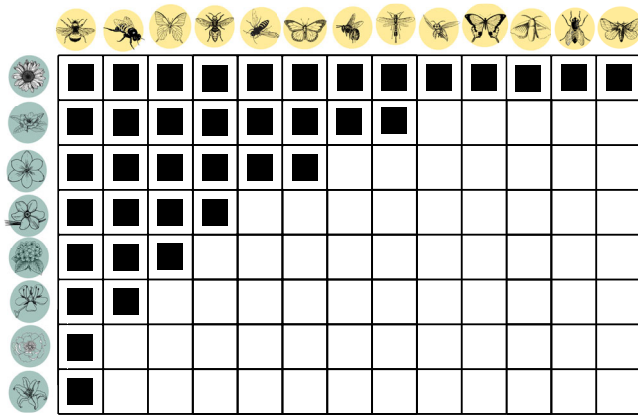
FIG. 1. Example of a perfectly nested biadjacency matrix for a bipartite plant-pollinator network. Each node represents a different species, and black boxes represent a mutualistic interaction between a plant and a pollinator. It can be observed that the interactions of any species form part of a subset of the interactions of all species with larger degree.

and hence are in turn connected to a variety of neighbors, including other generalists. This gives to nested biadjacency matrices its distinctive triangular shape composed by a robust core of connections among generalists to which specialists cling (see Fig. 1). Although real ecosystems are never maximally nested, numerous examples of mutualistic communities across the globe have been found to be noticeably nested, regardless of their differences in species composition, climate, or geographical location. Not only does it seem to be a rather universal pattern in ecology, but nestedness has also been observed in other areas of study such as biogeography [2] and socioeconomy, particularly, for instance, in manufacturer-contractor networks [3], the world trade web [4,5], or the seller-buyer network of a fish market [6].

These outstanding observations have triggered an intense research aiming to properly define and measure nestedness [7–13] as well as explain its origin [14–17]. Indeed, the interest in deciding whether a system is nested and/or in determining its degree of nestedness goes beyond characterizing it from a merely topological viewpoint. In fact, it has been suggested that nestedness plays an important role in biodiversity persistence [18–20], a claim which is nevertheless the subject of an ongoing and intense debate [21–25]. Admittedly, disclosing how the structure of mutualistic communities, particularly nestedness, affects its inner dynamics is key to understanding the history and predicting the future of worldwide ecosystems, all the more since mutualistic relationships pervade nature and carry out essential services, from pollination to nitrogen fixation. While antagonistic interactions, like parasitism or predation, are certainly ecologically relevant as well, whether the resulting empirical networks are nested or not is a controversial issue [26,27], not to mention that theoretical studies have claimed that nestedness might be a destabilizing

configuration for antagonistic networks [19]. In addition, empirical data on how real communities map into more general networks that include simultaneously different types of interactions have just begun to appear lately [28,29]. Altogether, this explains why mutualistic networks, yet if a partial representation of the complexity of natural interactions, are consistently used as model ecosystems.

Recently, the pertinence of nestedness as a suitable indicator to characterize the dynamics of mutualistic communities has been challenged by various works. In fact, it has been argued that either nestedness has no significant impact on the coexistence of mutualistic communities [23] or it detrimentally affects local stability [22,30]. Moreover, other properties of the observed networks have been claimed to be significant drivers of the community dynamics [31,32]. In particular, the networks' degree assortativity or the degree heterogeneity have been identified as determinants of biodiversity persistence.

Interestingly, within this debate, it was argued that the results found by James *et al.* [31] were an artifact of the applied randomization process, which did not control for the degree sequences [33]. This claim was in turn answered by the authors of Ref. [34], who showed that their results still hold when randomizing the networks preserving strictly the degree sequences. This controversy leads to the key question of whether nestedness conceived as a global trait of the emerging architecture is actually a relevant and independent property, or contrarily, if it just derives from lower-order features of the interaction network. Distinguishing between genuine and redundant patterns is critical if one aims to obtain, from the present structure of, e.g., ecological communities, information about its past history and assembling. Therefore, the main goal of our work is, precisely, to understand whether nestedness is an independent pattern, or, instead, if it is naturally caused by other properties of the network. Importantly, answering this question would also allow us to solve another open challenge: Namely, what is, in this context, the right *null model* against which one should assess nestedness? The latter question is a relevant issue by itself [35,36], as any attempt to quantify nestedness implicitly involves the comparison with a null hypothesis.

Earlier investigations on the structural determinants of nestedness were concurrent on the relevance of the degree sequences. First, Medan *et al.* [37] theoretically showed that the geometric curve that delimits the region with interactions in an ideally nested matrix [7] can be ultimately related, in the continuous limit approximation, to the degree distributions of both guilds of the corresponding bipartite network. On the other hand, Joppa *et al.* [38] identified the degree sequences as a feature that considerably explains empirical nestedness, although they still claimed to find "a statistically significant excess of networks with unusual nestedness patterns." Finally, Johnson *et al.* [39] explored the emergence of correlations in a finite-size configuration

model and argued that degree heterogeneity together with dissassortativity are two crucial determinants of nestedness.

Furthermore, in the context of building null models to measure the significance of nestedness of real systems, other works have been confronted with the relationship between nestedness and heterogeneity of degrees, yet without fundamentally addressing it. A well-known null model that relates to our work is the *fixed-fixed* (FF) model, in which the degree sequences are strictly kept. Typically, its implementation consists of algorithmically randomizing the real network, producing null networks by swapping pairs of interactions in such a way that the degree sequences are not altered [40]. A common caveat of this model is that the number of null networks compatible with the constrained degree sequences might be highly limited. For instance, Ulrich *et al.* [12,35] observed poorly significant nested patterns when using the FF null model. However, they explained such a result arguing that the FF null model induces a bias in the sampling due to the fact that the generated null matrices closely resemble the real network, contrary to the case where a randomization that relaxes the degree sequences is applied. In their words [35], "This similarity makes it more difficult for the FF algorithm to detect nestedness." Similarly, Staniczencko *et al.* [30] pointed out the limitations of the FF model in assessing the nestedness significance, given that the number of possible null networks decreases as the nestedness of the real network increases.

Despite all these different hints pointing to the crucial question of the relevance of nestedness as an independent property, the debate still remains open [21]. The difficulty in obtaining a definitive answer to this problem is related to the nature of the methodology used in previous studies, which discussed the statistical correlations between observed properties of the real systems and the corresponding nestedness measures provided by the considered null models. Since these null models typically randomize *algorithmically* the observed network under some chosen constraint, they are often affected by correlations that are enhanced by the finite size of the studied networks, eventually introducing undesired statistical bias.

In order to address the aforementioned questions, here we adopt a completely different path. We derive mathematical expressions for the first two moments of the nestedness distribution, which depend only on the entries of the adjacency matrix. Contrary to the former studies discussed above, we construct a maximum entropy ensemble of networks where constraints are not enforced algorithmically but by maximizing its likelihood of appearance. This prevents the statistical bias that was present in the previous methodological approaches. Next, we demonstrate that the nestedness is determined by the degree sequences of the two branches of the corresponding bipartite graph, and therefore, that it is not an irreducible pattern. We validate our claim by showing that for each one

of the 167 mutualistic networks of an empirical set formed by diverse types of plant-animal mutualistic communities, the observed amount of nestedness in real ecosystems could solely arise from the empirical degree sequences. Our work also provides a null model that can be used to discriminate whether a real network shows a significant nested pattern or not. Additionally, we discuss the rationale of our findings and provide an explanation of our results in terms of the heterogeneity of the degree distributions and their entropic character. We round off the paper discussing the implications of our work. In particular, the fact that the astounding universality of nested patterns can be reinterpreted as a result of the distribution of the number of contacts indicates that evolution has not selectively favored a nested assemblage of mutualistic communities but simply a significant diversity in the degree of specialization in mutualistic relationships. More generally, our results are a reminder of the importance of systematically testing the significance of any observed network pattern, even more if one aims to study the influence of such a pattern on the network dynamics. These conclusions, although more directly impactful in the area of ecological and mutualistic systems, could also be relevant for our general understanding of the topology of networks and their role in social, biological, and technological systems.

II. CONSTRUCTION OF THE STATISTICAL ENSEMBLE

We first construct a maximum entropy statistical ensemble of networks for each of the empirical networks that we use to validate our theoretical insights under the constraint that the degree sequences in the ensemble match on *average* the empirical ones—this being true for the two guilds of the corresponding bipartite graph. This methodological approach has the advantage that possible missing links or overrated interactions, which might lead to impoverished ecological data [41], are dealt with in a proper way. Indeed, enforcing the randomized degree sequences to be equivalent to the empirical ones only on average limits the possible impact of having noise in the data, while assuring that the results are not dependent on specific details.

To construct the ensemble, we adopt a statistical physics perspective [42] by providing the probability of the appearance of each network in the ensemble. This varies from previous works which proposed randomizing the real network, although imposing similar conditions on the degrees, *algorithmically* [35,38,40]. Instead, here we apply a recently introduced randomizing scheme [43,44] and construct 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 a constraint. Yet, it is not fully determined due to the presence of free Lagrange multipliers resulting from the constrained optimization of the entropy. Following Squartini and Garlaschelli [43] and
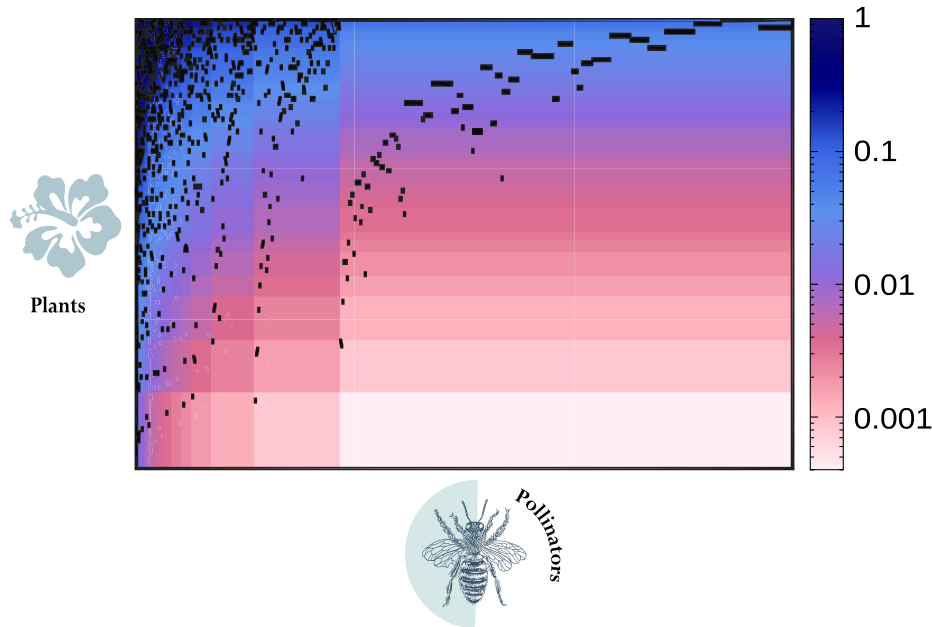
FIG. 2. Comparison between empirical mutualistic interactions and probability of interacting in the ensemble. The probability of interaction between species in the statistical ensemble given by $\langle \mathbf{B}^* \rangle$ is shown as a color heat map for the plant-pollinator network recorded by Inoue *et al.* [46]. The empirical corresponding bipartite matrix of interactions $\mathbf{B}^*$ is superimposed in black. Both plant and pollinator species are 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.

Garlaschelli and Loffredo [45], we impose that the degree sequences of the empirical network have to be found with maximum likelihood (see the Appendix A for details). This enforces, as aimed, the ensemble's mean degree sequences to match the empirical ones, whilst precluding common biases of other sampling techniques [45].

Eventually, this maximization of the likelihood provides a set of coupled equations to solve for the Lagrange multipliers [in principle, one equation per node; see Eqs. (A7) and (A8) in Appendix A]. Hence, determining the statistical random ensembles of our 167 empirical networks entails solving computationally 167 optimization problems. For each network, we numerically find the Lagrange multipliers that maximize the likelihood using two different independent algorithms: (i) simulated annealing, which is a global pseudorandom numerical method for optimizing the likelihood and (ii) a deterministic gradient-based algorithm for solving nonlinear systems of equations (see Appendix B for more information on the numerical implementation).

A primary advantage of constructing a maximum-likelihood and maximum entropy ensemble is that in the case of local constraints (like the degree sequences), the probability of the existence of a graph in the ensemble can be exactly factorized into the probabilities of the existence of a link between any two species [43]. In this way, we can obtain the average probability that two potential mutualistic partners interact in the statistical ensemble. If we call a real bipartite matrix $\mathbf{B}^*$, then this probability matrix will be represented by $\langle \mathbf{B}^* \rangle$. Therefore, after numerically determining each optimal set of Lagrange multipliers, we build the matrix containing the average probability of interaction corresponding to each empirical network (see an example in Fig. 2).

III. STATISTICAL MEASURES OF NESTEDNESS

We perform the statistical measures on the ensemble following one of the two following approaches. On the one hand, as long as the property that we aim to evaluate can be analytically formulated in terms of the elements of the bipartite adjacency matrix, Squartini and Garlaschelli showed [43] 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. In other words, it is not necessary to sample the ensemble; instead, the mean and the standard deviation of the nestedness index can be analytically calculated. On the other hand, one can always sample the ensemble in order to study the statistics of the target property on a generated unbiased sampling [47]. In this paper, we focus on the analytical solution to the debate about the origin of nestedness. Nevertheless, we check that our results also hold when using the sampling approach.

In this section, we start by describing how to obtain the analytical expressions of the first two moments of a network property in the ensemble. Second, we apply them to derive the analytical expressions for the mean and the standard deviation of the two indices for nestedness for which this is possible, as we discuss in the following subsections: the well-known nestedness metric based on overlap and decreasing fill (hereafter, NODF) [9] and the recently proposed spectral radius [30].

A. General analytical expressions of the first two moments of the distribution of a given property of the network

Here, we describe how to obtain statistical measures in the random ensemble through analytical expressions. Let us call a property X and its average across the statistical ensemble $\langle X \rangle^*$. The asterisk superscript indicates that the statistical ensemble is built for a given real bipartite matrix $\mathbf{B}^*$. When the property X can be calculated as an analytic function of the bipartite matrix $\mathbf{B}$, then it is possible to derive an analytical expression of the first and second moments of X in terms of the average probabilities of interaction $\langle \mathbf{B}^* \rangle$. Following Squartini and Garlaschelli [43], for the bipartite case, these expressions read

$$\langle X \rangle^* \simeq X(\langle \mathbf{B} \rangle^*), \tag{3}$$

$$\sigma_X \simeq \sqrt{\sum_{p=1}^{N_P} \sum_{a=1}^{N_A} \left(\frac{\partial X(\mathbf{B})}{\partial b_{pa}} \right) \Bigg|_{\mathbf{B}=\langle \mathbf{B} \rangle^*}^2 \sigma_{b_{pa}}^2}, \tag{4}$$

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 needs to be Gaussian distributed in the random ensemble.

B. Derivation of the analytical expressions for NODF

In this subsection, we present the analytical expressions for the average and the standard deviation of nestedness measured by the index known as NODF. We choose this metric among the vast range of indices in the literature due to a variety of reasons: First, it can be calculated through an analytical and compact expression in terms of the matrix elements. Second, and contrary to other metrics, its definition is based on a clear and explicit quantification of conditions in Eqs. (1) and (2), precluding any type of geometric or algorithmic approach. Finally, it is widely used not only in ecology but also in network applications to economics [6,44] or sociology [48].

1. Definition and distribution of NODF

The NODF index involves two contributing factors to nestedness: decreasing fill, that quantifies to what extent, after ordering the rows and columns of the matrix, the

degree sequences strictly decrease, and paired overlap, that accounts for the number of shared partners between all pairs of columns (rows), normalized by the smaller degree. By gathering the sequential analysis indicated by Almeida-Neto et al., [9] we propose the following compact expression to calculate the NODF:

$$\begin{aligned}
\text{NODF}(\mathbf{B}) = &\frac{1}{K} \sum_{i<j}^{N_P} \left\{ [1 - \theta(v_j - v_i)] \frac{\sum_{a=1}^{N_A} b_{ia} b_{ja}}{v_j} \right\} \\
&+ \frac{1}{K} \sum_{k<l}^{N_A} \left\{ [1 - \theta(h_l - h_k)] \frac{\sum_{p=1}^{N_P} b_{pk} b_{pl}}{h_l} \right\},
\end{aligned} \tag{5}$$

$$\text{where } K = \frac{N_P(N_P - 1) + N_A(N_A - 1)}{200}. \tag{6}$$

Here we use the following notation: v_p is the degree of plant p and h_a the degree of animal a. The double sums run over two indices, and we consider that the bipartite adjacency matrix is labeled as shown in Fig. 3, such that row i is placed above row j and column k at the left of column l. The K factor contains the normalization over the number of all possible pairs and the fact that the NODF is defined to take values between 0 and 100. Finally, the θ stands for the Heaviside step function, which is zero when its argument is negative, and 1 if its argument is positive or zero. As a result, the $1 - \theta(v_j - v_i)$ term encapsulates the decreasing fill condition.
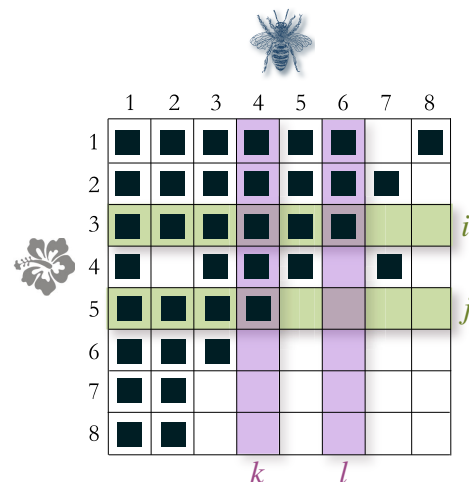

FIG. 3. Example of an ordered matrix of interactions not perfectly nested. Species of both guilds are ordered in decreasing degree, and the numbered labels indicate their rank (the larger the degree, the smaller the rank). The indexes i, j, k, and l illustrate our notation for rows and columns.

In fact, from now on we use the following abbreviations for the decreasing fill term:

$$DF_{ij} = 1 - \theta(v_j - v_i), \text{ such that,}$$

$$\text{if } v_j \geq v_i \quad \text{then } DF_{ij} = 0,$$

$$\text{and if } v_j < v_i \quad \text{then } DF_{ij} = 1, \tag{7}$$

$$DF_{kl} = 1 - \theta(h_l - h_k), \text{ such that,}$$

$$\text{if } h_l \geq h_k \quad \text{then } DF_{kl} = 0,$$

$$\text{and if } h_l < h_k \quad \text{then } DF_{kl} = 1. \tag{8}$$

Despite being a popular metric, some authors have raised concerns about the use of the NODF to measure nestedness. In particular, Staniczenko et al. [30] showed that the NODF is unable to detect nested patterns when the proportion of repeated degrees in the network is large. This shortcoming is due to the *decreasing fill* factor, which heavily penalizes degree degeneracy. In order to ensure that our results are not affected by this limitation in the sensibility of the NODF, we obtain as well the analytical expressions of the first two moments of an alternative version of the metric called the *stable NODF* proposed by Mariani et al. [49]. This metric is more robust against slight variations in the degree sequence and, importantly, solves the drawbacks outlined by Staniczenko et al. [30] since it does not include the decreasing fill term. The definition of the stable NODF as well as the analytical expressions of the first two moments of its distribution in the ensemble can be found in Appendix C. Still, we keep our main focus on the NODF given its widespread use for measuring nestedness, together
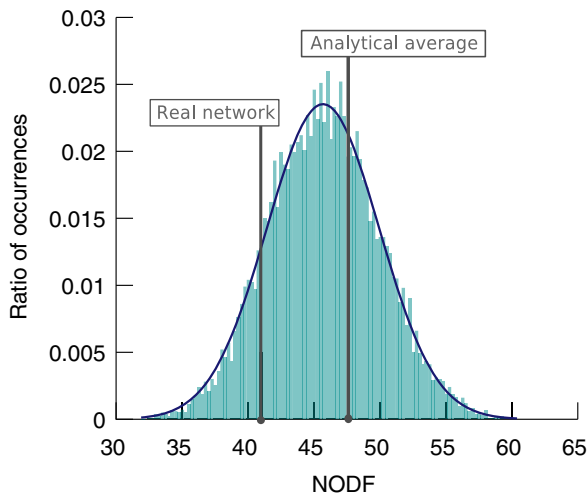

FIG. 4. Nestedness distribution for a sampling of the statistical ensemble corresponding to the empirical networks by Small [50]. In blue, the fit of a Gaussian function using the mean and standard deviation extracted from the distribution (mean $\mu = 45.8$ and standard deviation $\sigma = 4.2$). In gray, the values of the nestedness of the real network and of the analytical average, which correspond to the average computed using the analytical expression in Eq. (9).

with the fact that the mentioned lack of sensibility is not particularly relevant but for small or highly dense networks.

We next verify that the NODF is Gaussian distributed in the ensemble, as required if we aim to apply Eqs. (3) and (4) by performing a check on a subset of the empirical networks. To this end, for each of the corresponding statistical ensembles, we generate a sample of 10^4 networks obeying the probability of link existence given by $\langle \mathbf{B} \rangle^*$. We then compute the nestedness of each sampled network using the NODF in order to generate the nestedness distribution. In all cases, we successfully fit a Gaussian function (see Fig. 4 for an example).

2. Analytical expression for the first moment of the NODF

The analytical and packed expression for the NODF that appears in Eq. (5) can then be plugged into Eq. (3). Accordingly, we obtain that the first moment of the randomized NODF for a given real bipartite matrix $\mathbf{B}^*$ reads

$$\langle \text{NODF}(\mathbf{B}) \rangle^* = \frac{1}{K} \sum_{i<j}^{N_P} \left\{ DF_{ij} \frac{\sum_{a=1}^{N_A} \langle b_{ia} \rangle \langle b_{ja} \rangle}{\sum_{a=1}^{N_A} \langle b_{ja} \rangle} \right\}$$
$$+ \frac{1}{K} \sum_{k<l}^{N_A} \left\{ DF_{kl} \frac{\sum_{p=1}^{N_P} \langle b_{pk} \rangle \langle b_{pl} \rangle}{\sum_{p=1}^{N_P} \langle b_{pl} \rangle} \right\}. \tag{9}$$

Note 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 constraints. Additionally, this condition warrants that the ordering of the matrix be equal to the original one, which is important since the NODF is ordering dependent through the decreasing fill terms. It is also of interest to remark that the previous expression can be understood in probabilistic terms. Indeed, given that $\langle b_{pa} \rangle = p_{pa}$, 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 the pair of animal species k and l, the overlap term results in

$$\frac{\sum_{p=1}^{N_P} \langle b_{pk} \rangle \langle b_{pl} \rangle}{\sum_{p=1}^{N_P} \langle b_{pl} \rangle} = \frac{\sum_{p=1}^{N_P} p_{pk} p_{pl}}{\sum_{p=1}^{N_P} p_{pl}}. \tag{10}$$

3. Analytical expression for the second moment of the NODF

The standard deviation is given by Eq. (4), which for the NODF reads

$$\sigma_{\text{NODF}} = \sqrt{\sum_{p=1}^{N_P} \sum_{a=1}^{N_A} \left(\frac{\partial \text{NODF}(\mathbf{B})}{\partial b_{pa}} \right) \Big|_{\mathbf{B} = \langle \mathbf{B} \rangle^*}^2 \sigma_{b_{pa}}^2}$$

$$\text{with} \quad \sigma_{b_{pa}}^2 = p_{pa}(1 - p_{pa}), \tag{11}$$

where we use the fact that the existence of a link in the network is a Bernoulli process. Furthermore, the derivative with respect to a general matrix element b_{rc} (the index r stands for rows and c stands for columns) can be split into the contributions of plants and of animals:

$$\frac{\partial \text{NODF}(\mathbf{B})}{\partial b_{rc}} = \frac{\partial \text{NODF}(\mathbf{B})_{\text{plants}}}{\partial b_{rc}} + \frac{\partial \text{NODF}(\mathbf{B})_{\text{animals}}}{\partial b_{rc}}. \tag{12}$$

After deriving, we obtained that

$$K\frac{\partial \text{NODF}(\mathbf{B})_{\text{plants}}}{\partial b_{rc}} = \sum_{j=r+1}^{N_P} DF_{rj}\frac{b_{jc}}{v_j} + \sum_{i=1}^{r-1} DF_{ir}\frac{b_{ic}}{v_r}$$
$$- \sum_{i=1}^{r-1}\sum_{a=1}^{N_A} DF_{ir}\frac{b_{ia}b_{ra}}{v_r^2}, \tag{13}$$

$$K\frac{\partial \text{NODF}(\mathbf{B})_{\text{animals}}}{\partial b_{rc}} = \sum_{l=c+1}^{N_A} DF_{cl}\frac{b_{rl}}{h_l} + \sum_{k=1}^{c-1} DF_{kc}\frac{b_{rk}}{h_c}$$
$$- \sum_{k=1}^{c-1}\sum_{p=1}^{N_P} DF_{kc}\frac{b_{pk}b_{pc}}{h_c^2}, \tag{14}$$

which after being plugged into Eq. (11) provides an analytical expression for the standard deviation of the distribution of the NODF in the ensemble.

C. Derivation of the theoretical expressions for the spectral radius

In this subsection, we derive the theoretical expressions for calculating the average and standard deviation of nestedness using the so-called *spectral radius* [30]. We perform the statistical measures using this metric due to its increasing popularity among nestedness indices, and, moreover, due to its computational advantages: First, it is a mathematical property of the graph which does not depend on the ordering of the matrix, and second, its numerical calculation is fast. Nonetheless, in this subsection we argue as well that the measures obtained with the spectral radius should be handled with care, since this metric does not reliably quantify nestedness at a fine scale and, moreover, it is not normalized.

1. Definition and distribution of the spectral radius

The spectral radius was recently proposed by Staniczenko et al. [30] as an alternative metric for nestedness that directly relies on the spectral properties of the adjacency matrix. Let us call $\mathbf{I}$ the identity matrix and $\mathbf{A}$ the adjacency matrix of a bipartite matrix $\mathbf{B}$, such that

$$\mathbf{A} = \begin{pmatrix} 0 & \mathbf{B} \\ \mathbf{B}^{\intercal} & 0 \end{pmatrix}, \tag{15}$$

which is a square, symmetric, and non-negative matrix, given that $a_{i,j} \geq 0$. The spectral radius of the matrix $\mathbf{A}$ (also called the *dominant eigenvalue* or *largest eigenvalue*) is defined as follows:

$$\rho(\mathbf{A}) = \max\{|\lambda_i|\}, \tag{16}$$

where λ_i for $i \in \{1, \ldots, n\}$ are the eigenvalues of $\mathbf{A}$; thus, the roots of the equation $\det(\mathbf{I}\lambda - \mathbf{A}) = 0$. Since $\mathbf{A}$ is a symmetric matrix, $\lambda_i \in \Re \ \forall \ i$.

The capability of the spectral radius for quantifying the degree of nestedness of a network is rooted in a theorem by Bell et al. [51] that states that within the set of networks having the same number of links and nodes, the one yielding the maximum spectral radius will be perfectly nested. In fact, Staniczenko et al. [30] showed that more nested networks tend to have larger spectral radius. However, importantly, this relation is only true in statistical terms. Indeed, the results of Staniczenko et al. reveal that, if we take two slightly different networks, the one with the largest spectral radius is *not* necessarily the most nested (see, for instance, Fig. 1 from Ref. [30]). Therefore, the sensibility of the spectral radius at a fine scale (i.e., to distinguish between small differences in the degree of nestedness of two networks) is rather limited. Another caveat of the spectral radius is that it is not normalized. This implies that nestedness measures are affected by network properties like the density of links or the size,
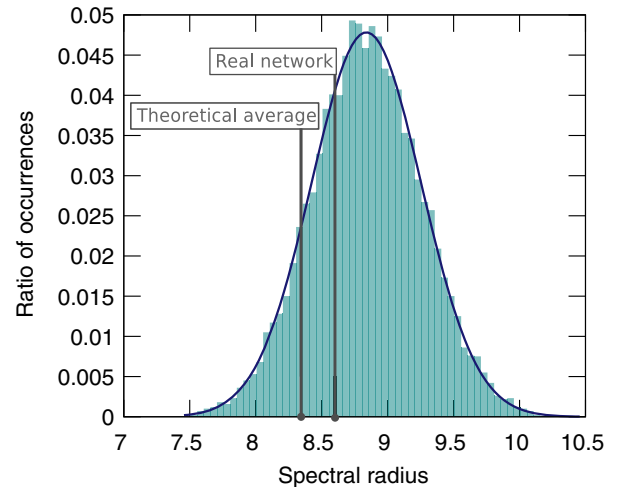

FIG. 5. Distribution of the spectral radius over the ensemble calculated for the real network collected by Small [50] for a sampling made of 10^4 networks. In blue, the fit of a Gaussian function using the mean and standard deviation extracted from the distribution (mean $\mu = 8.8$ and standard deviation $\sigma = 0.4$). In gray, the values of the nestedness of the real network and of the theoretical average, which correspond to the average computed using the theoretical expression in Eq. (19).

thus hindering the comparison among networks which do not share those characteristics.

Next, we check whether the spectral radius is Gaussian distributed over the ensemble. In particular, for each network in our dataset, we generate a sample of 10^4 networks obeying the probability of link existence provided by $\langle \mathbf{B} \rangle^*$. Then, we calculate the spectral radius of each sampled network algorithmically using the R package RARPACK [52]. Finally, we verify that the resulting distribution is indeed normal, as can be seen in Fig. 5.

2. Theoretical expression for the first moment of the spectral radius

Given that $\rho(\mathbf{A})$ is a function of the matrix entries of $\mathbf{A}$, we can apply the linear approximation proposed by Squartini and Garlaschelli in Eq. (3) in order to estimate the average over the ensemble computed for a real bipartite matrix $\mathbf{B}^*$:

$$\langle \rho(\mathbf{A}) \rangle^* \approx \rho(\langle \mathbf{A} \rangle^*), \tag{17}$$

where

$$\langle \mathbf{A} \rangle^* = \begin{pmatrix} 0 & \langle \mathbf{B} \rangle^* \\ \langle \mathbf{B} \rangle^{*\mathsf{T}} & 0 \end{pmatrix}. \tag{18}$$

These expressions mean that the average spectral radius can be found as

$$\rho(\langle \mathbf{A} \rangle^*) = \max\{|\langle \lambda_i \rangle|\}, \tag{19}$$

where $\langle \lambda_i \rangle$ are the roots of the equation

$$\det(\mathbf{I}\langle \lambda \rangle - \langle \mathbf{A} \rangle^*) = 0. \tag{20}$$

In practice, Eq. (20) has to be solved numerically, which implies that no analytical expression for the average of the spectral radius exists. In particular, we numerically implement the calculation of the spectral radius of each matrix $\langle \mathbf{A} \rangle^*$ using the R package RARPACK [52].

3. Analytical expression for the second moment of the spectral radius

Using Eq. (4), the standard deviation of the spectral radius over the ensemble can be estimated by

$$\sigma_\rho \simeq \sqrt{\sum_{p=1}^{N_P} \sum_{a=1}^{N_A} \left(\frac{\partial \rho(\mathbf{A})}{\partial A_{pa}} \right)\Big|^2_{\mathbf{A}=\langle \mathbf{A} \rangle^*} \sigma^2_{A_{pa}}}. \tag{21}$$

Here, the calculation of the derivative of the spectral radius $\{[\partial \rho(\mathbf{A})]/\partial A_{pa}\}|_{\mathbf{A}=\langle \mathbf{A} \rangle^*}$ is nontrivial given that there is no general analytical expression for the spectral radius because Eq. (20) needs to be solved numerically. Nonetheless, we now show how is it possible to obtain such a derivative by applying the results by Deutsch and Neumann [53].

Let us start by assuming that $\mathbf{M}$ is a square, nonnegative, and irreducible matrix. Then, it is known that its spectral radius $\rho(\mathbf{M})$ is a simple eigenvalue, and it is equal to its Perron root. Since $\rho(\mathbf{M})$ is a simple eigenvalue, it has multiplicity 1, and it is possible to obtain its first derivatives with respect to $\mathbf{M}_{ij}$. Indeed, if we denote by $\mathbf{D}$ the matrix whose matrix elements are

$$\mathbf{D}_{ij} = \frac{\partial \rho(\mathbf{M})}{\partial \mathbf{M}_{ij}}, \tag{22}$$

then following Ref. [53], $\mathbf{D}$ can be computed using the expression

$$\mathbf{D} = (\mathbf{I} - \mathbf{Q}\mathbf{Q}^\#)^\mathsf{T}. \tag{23}$$

Here, $\mathbf{Q}$ is a special type of matrix known as an *M matrix* [54] and defined as

$$\mathbf{Q} = \rho(\mathbf{M})\mathbf{I} - \mathbf{M}, \tag{24}$$

while $\mathbf{Q}^\#$ is the *group inversion* of $\mathbf{Q}$ [55]. The group inversion is a more general type of inverse that can be applied as well to singular matrices. For certain types of matrices, the group inverse is equivalent to another class of inversion known as the *Moore-Penrose inverse*. This is true if and only if the matrix of study is *range Hermitian* [55]. One of the conditions that warrants that a matrix is range Hermitian is the following:

$$\text{range}(\mathbf{Q}) = \text{range}(\mathbf{Q}^H), \tag{25}$$

where $\mathbf{Q}^H$ is the conjugate transpose (also called Hermitian conjugate) of $\mathbf{Q}$. If we now assume that $\mathbf{Q}$ is range Hermitian, Eq. (23) can be rewritten as

$$\mathbf{D} = (\mathbf{I} - \mathbf{Q}\mathbf{Q}^\dagger)^\mathsf{T}, \tag{26}$$

where $\mathbf{Q}^\dagger$ represents the Moore-Penrose inverse of $\mathbf{Q}$.

Let us show now that Eq. (26) can be used to calculate the derivative with respect to our matrix of interest $\mathbf{A}$, in particular, in the case where $\mathbf{A} = \langle \mathbf{A} \rangle^*$. First, we show that $\langle \mathbf{A} \rangle^*$ fulfills the conditions that allow us to apply Eq. (23). Next, we prove that a matrix $\mathbf{Q}_A$ defined as

$$\mathbf{Q}_A = \rho(\langle \mathbf{A} \rangle^*)\mathbf{I} - \langle \mathbf{A} \rangle^* \tag{27}$$

is a range-Hermitian matrix and, consequently, $\mathbf{Q}_A^\# = \mathbf{Q}_A^\dagger$.

First, we know already that $\langle \mathbf{A} \rangle^*$ is a square and nonnegative matrix, yet it remains to be shown whether it is irreducible. A matrix is said to be irreducible if and only if its corresponding graph is strongly connected; that is, if it is possible to find a path that connects any pair of nodes of the network. Because $\langle \mathbf{B} \rangle^*$ is a complete bipartite

graph ($\langle b_{ij} \rangle^* > 0 \ \forall \ i, j$), then it is clear that it is strongly connected, and therefore, $\langle \mathbf{A} \rangle^*$ is an irreducible matrix.

Second, the condition for the matrix $\mathbf{Q}_A$ to be range Hermitian is provided by Eq. (25). In our case, $\rho(\mathbf{A}) \in \mathfrak{R}$ and $\mathbf{A}_{ij} \in \mathfrak{R} \ \forall \ i, j$, therefore, $\mathbf{Q}_{A,ij} \in \mathfrak{R} \ \forall \ i, j$. From this, it follows that

$$\text{range}(\mathbf{Q}_A^H) = \text{range}(\mathbf{Q}_A^\mathsf{T}), \tag{28}$$

where we use that the conjugate transpose of a real matrix is simply its transpose. We still need to prove that

$$\text{range}(\mathbf{Q}_A^\mathsf{T}) = \text{range}(\mathbf{Q}_A). \tag{29}$$

This condition is equivalent to

$$\text{row space of } \mathbf{Q}_A = \text{column space of } \mathbf{Q}_A. \tag{30}$$

Note that this is not true in general. In our case, given that $\mathbf{Q}_A$ is a square and symmetric matrix, its row and column spaces are equal, and therefore, Eqs. (29) and (30) are verified. This proves that $\mathbf{Q}_A$ is range Hermitian, and consequently, its group inverse is equivalent to its

Moore-Penrose inverse. With this proof, we show that a matrix $\mathbf{D}_A$ defined by

$$\mathbf{D}_{A,ij} = \frac{\partial \rho(\mathbf{A})}{\partial \mathbf{A}_{ij}} \bigg|_{\mathbf{A} = \langle \mathbf{A} \rangle^*} \tag{31}$$

can be computed using the following expression:

$$\mathbf{D}_A = (\mathbf{I} - \mathbf{Q}_A \mathbf{Q}_A^\dagger)^\mathsf{T}. \tag{32}$$

We implement Eq. (32) using the R MASS package [56], in particular, the function ginv to calculate the Moore-Penrose inverse.

IV. SIGNIFICANCE OF EMPIRICAL NESTED PATTERNS

For each of the 167 empirical networks, we obtain the matrix giving the probability of interaction between any two species of different guilds. This set includes three different kinds of mutualistic communities: plant-pollinator, seed-disperser and plant-ant (see Appendix J). Using Eqs. (9) and (11), we are able to compute the
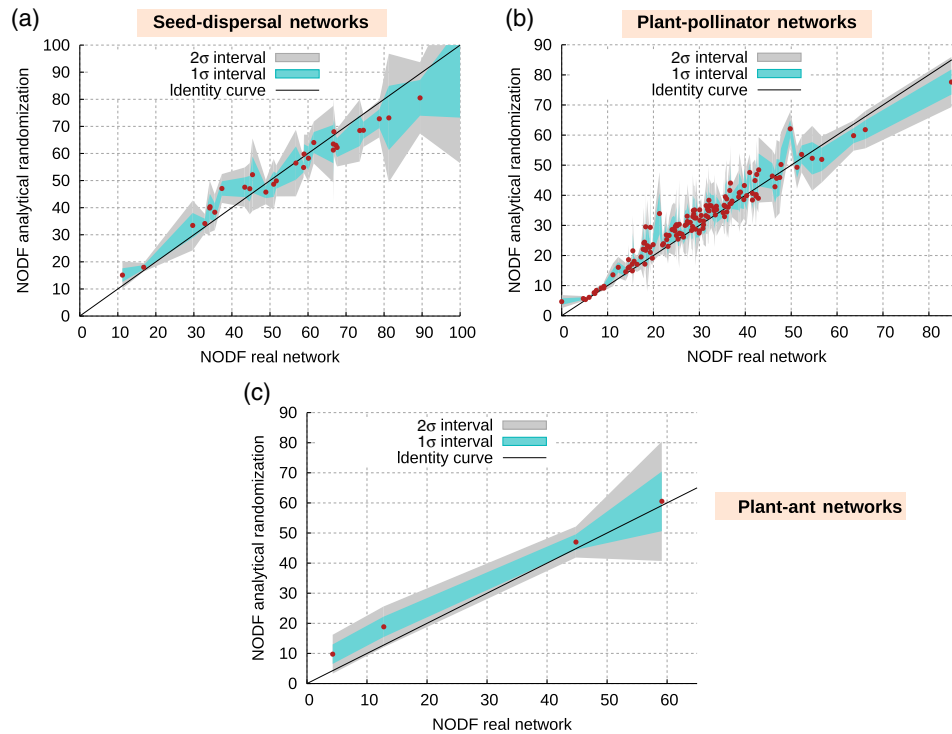

FIG. 6. Significance of the nestedness of real networks. The figure shows the empirical measure of nestedness against the average value of nestedness in the generated statistical ensemble (red dots) for the 167 real mutualistic networks. The three panels correspond to different kinds of mutualistic systems: (a) seed-dispersal, (b) plant-pollinator, and (c) plant-ant. In particular, the ant-plant mutualism is where ants provide plants with services like defense from herbivores, pollination, and/or gathering of nutrients, while plants offer shelter and/or food. The shadowed areas represent 1 (teal) and 2 (light gray) standard deviations of the mean. Further details about the number of networks whose nestedness are within these boundaries are provided in Table I. A detailed significant test results in only three networks having a statistical significant (in all cases under-represented) nestedness value. Overall, the results indicate that the nestedness of these mutualistic networks is not significant.

TABLE I. Results disentangled into communities showing the fraction of networks (abbreviated above as "ntws") whose discrepancy between the real and randomized nestedness is less than or equal to 1σ or 2σ. Nestedness is measured with the NODF.

Type of community	Fraction of ntws with $|z\ \text{score}| \leq 1$		Fraction of ntws with $|z\ \text{score}| \leq 2$	
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%

expectation value of nestedness measured by the NODF, $\langle \text{NODF}(\mathbf{B}) \rangle^*$, and its standard deviation, for each empirical network in our dataset. A comparison between the expected value of nestedness calculated over the statistical ensemble corresponding to each real network and the actual nestedness of the real network shows a striking agreement; see Fig. 6. As reported in Table I, the absolute difference between these two quantities is less than 1 standard deviation for 100 out of 167 networks (59.9%), rising to 158 out of 167 networks (94.6%) if we account for 2 standard deviations. After performing a multiple testing correction (see Appendix D), we find that only three out of the 167 empirical networks show significant nestedness (corrected p value < 0.01). The three of them, which are of relatively small size (≤ 55 species), are found to be *less* nested than predicted by the statistical ensemble.

Moreover, we verify that our results are not affected by the shortcomings related to the decreasing fill factor by repeating the above calculations using the index stable NODF. In fact, our conclusions not only hold but are also strengthened when using this metric, since we find 118 out of 167 networks (70.7%) within a distance of 1 standard deviation to the mean, and 162 out of 167 networks (97.0%) within 2 standard deviations (see Appendix E).

Additionally, in order to ensure that our findings are not an artifact of using the NODF metric, we perform the same
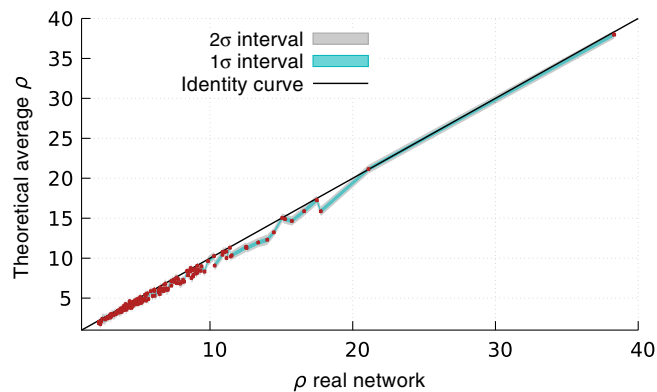
analysis for the spectral radius, as we explain in Sec. III C. As shown in Fig. 7 and Table II, this supplementary analysis produces results that are generally in agreement with those reported above for the NODF metric. After performing the multiple testing correction, we find that 22 out of 167 networks unexpectedly nested (corrected p value < 0.01). Although this proportion is larger than the one found by the NODF, we argue that, given the poor performance of the spectral radius metric at fine scales, it is not necessarily an indication of the significant nested patterns but a consequence of the intrinsic limitations of the metric.

Finally, we also verify that performing the measures by sampling the ensemble and then computing the distribution of nestedness gives compatible results, both for the NODF (see Appendix F) and the spectral radius (see Appendix G).

V. NETWORK PROPERTIES DETERMINING NESTEDNESS EMERGENCE

In light of the previous results, the next question to be considered is whether we can determine which characteristic of the degree sequences controls how nested a network is. Taking into account that the degree distributions of mutualistic communities have been reported to commonly follow a (truncated) power law [57], 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 make use of a self-organizing network model that is devised with the aim of optimizing the nestedness of a network [58] by rewiring existing links (see Appendix H). After applying this

FIG. 7. Significance of the nestedness of real networks measured with the spectral radius ρ. The figure shows the theoretical average value of nestedness in the statistical ensemble against the empirical measure of nestedness (red dots) for the 167 real mutualistic networks. The shadowed areas represent 1 (teal) and 2 (light gray) standard deviations of the mean.

TABLE II. Fraction of networks whose discrepancy between the real and randomized nestedness is less than or equal to 1σ or 2σ for nestedness measured by the spectral radius.

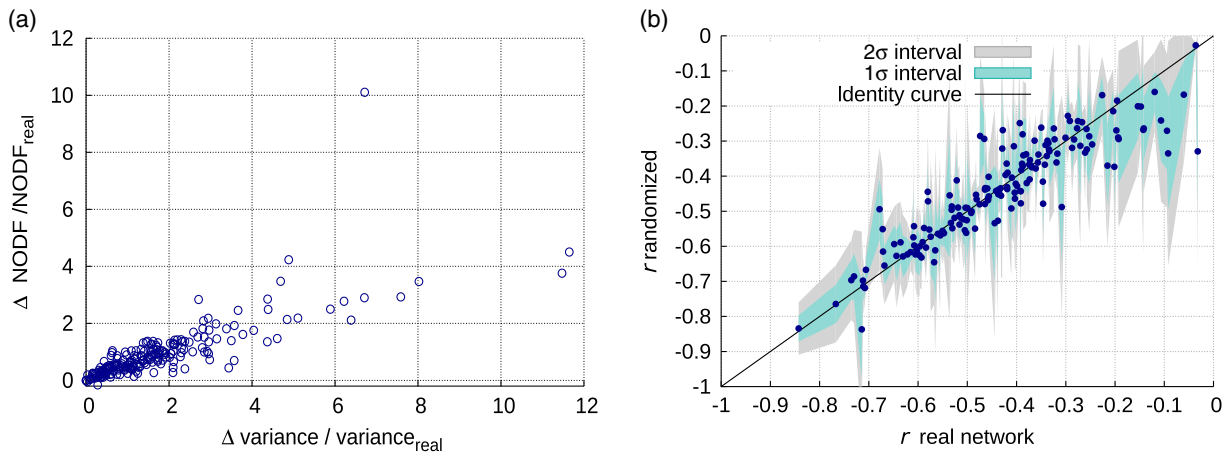| Fraction of ntws with $|z\ \text{score}| \leq 1$		Fraction of ntws with $|z\ \text{score}| \leq 2$	
67 out of 167	40.1%	114 out of 167	68.3%

FIG. 8. Determinants of nestedness. (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 use the rewiring algorithm described in Appendix H. 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 find 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. (b) 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 indicates that both real networks and the average of the randomized ensemble are naturally disassortative.

algorithm to our empirical set of networks, we find that the resulting degree sequences are, with respect to the original ones, more heterogeneous and that the final networks are more nested; see Fig. 8(a). This observation allows us to bridge the gap between two structural features that have been classically treated separately, although previous works have already suggested their connection [16,39]. Interestingly enough, the relationship between a 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 [18] or the network's structural stability [19,20], have been recently successfully associated with other properties such as the heterogeneity itself [32] or the species' degree [25,31].

Moreover, accounting for the heterogeneity offers some further insight into 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 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 supergeneralists as well as by finite-size effects, usually there is still room for reshuffling links (like in the "swapping algorithm" [40]). 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 (see Appendix I).

Why then does our algorithm generate disassortative networks as shown in Fig. 8(b)? Here, the particularity that we use a maximally entropic ensemble plays a crucial role. Johnson et al. [59] 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. To put it simply, this occurs 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. Johnson et al. [39] also observed this fundamental relationship among heterogeneity, disassortativity, and nestedness, yet using a finite-size configuration model. In our null model, on the other hand, finite-size correlations are decreased since the degree sequences are allowed to vary, thus showing that the emergence of nestedness is a genuine entropic consequence of degree heterogeneity.

VI. DISCUSSION

The findings above are of utmost importance in at least two fundamental aspects. First, they demonstrate that, given the degree sequence of real networks, the observed nestedness is not significant—at least for almost all networks here analyzed. Therefore, they show that nestedness is not an independent pattern, in sharp contrast to the widely extended belief that it is so. In other words, these results reveal that the observed nested structure of the

ecological communities studied is, in fact, a mere entropic consequence of the heterogeneity of the degree sequence. Second, regarding the long-standing controversies about the use of a proper null model for nested networks [35], our findings point out the need of incorporating the information contained in the degree sequences. Moreover, our results indicate that an appropriate null model is the set of exponential random graphs for which the probability of finding a graph having the empirical degree sequences is maximized [43]. This approach exhibits at least two main advantages with respect to previous null models. In the first place, it overcomes the finite-size effects of the FF model, which does not allow for a proper algorithmic randomization [12,35,40], hence restraining too narrowly the exploration of the phase space of null matrices. With respect to models that preserved the degrees on average by constructing a probability of interaction proportional to the species degrees [60], it has been shown that this form of the probability is subject to bias since it is not a maximum-likelihood choice [45]. In fact, while exponential random graph models have been extensively used in other fields like sociology [61] or economics [44], their possible application to ecological networks is still largely unexplored. Thus, we propose that the methodology implemented here (along with the theoretical expressions of the two first moments derived for the NODF index and the spectral radius) could be a general tool to asses nestedness' significance.

Concerning the question of the emergence of nestedness, 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 [18] or as a by-product of processes such as the assembling rules [62–64]. Our results imply that no selective pressure is required for nestedness to appear, which does not exclude, however, that such a pressure could have shaped the degree sequences. Even though such conclusions do not invalidate nestedness' usefulness as an indicator of stability or robustness, we like to underline that our findings clearly demonstrate that the degree sequences are the lower-order determinants of nestedness. Accordingly, the nested pattern is not *more* informative of the evolutionary history of real systems than their degree sequences alone.

Furthermore, it is important to recall that the networks of our study are often both spatially and temporally aggregated. Given the significant variability of species' interactions along time [65] and space [66], the fact that nestedness emerges from a local property like the degrees is a parsimonious explanation. Indeed, degree heterogeneity is a general feature, not only characteristic of the aggregated network but also of its spatial and temporal counterparts. On the other hand, if nestedness were an independent pattern, its emergence would require a specific selection of interactions across time and space, which is a much more intricate process and hence a less simple justification of the origin of nestedness.

In concluding, these results highlight the interest of focusing on the ecological and evolutionary mechanisms that have led to the coexistence of both specialized and generalized mutualisms in the same community [67,68], giving rise to the observed high heterogeneity of the degree sequences. 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.

ACKNOWLEDGMENTS

APPENDIX A: DETAILS ON THE MAXIMUM ENTROPY AND MAXIMUM-LIKELIHOOD ENSEMBLE

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

1. General randomizing scheme

We denote a network in the ensemble by its graph $\mathbf{G}$, and we call $\mathbf{G}^*$ the graph corresponding to the real network. We characterize the ensemble by the probability of the occurrence of each of its elements $P(\mathbf{G})$. Following Refs. [42,43], we determine the maximum entropy ensemble such that a set of constraints are satisfied on average, thus allowing slight mismatches across the ensemble. This methodology is equivalent to constructing a *grand-canonical ensemble* (opposite of the *microcanonical ensemble*, which corresponds to the FF null model, where the constraints need to be exactly met).

We start by maximizing the Shannon-Gibbs entropy defined as

$$S = -\sum_{\mathbf{G}} P(\mathbf{G}) \ln P(\mathbf{G}), \tag{A1}$$

where the sum runs over all the graphs G in the ensemble. This maximization leads to the exponential random graph model, which reads

$$P(G/\vec{\theta}) = \frac{e^{-H(G,\vec{\theta})}}{Z(\vec{\theta})}, \qquad (A2)$$

with H being the graph Hamiltonian such that $H(G,\vec{\theta}) = \vec{\theta} \cdot \vec{C}(G)$, and being Z the partition function $Z = \sum e^{-H(G)}$. The set of variables $\vec{\theta}$ are the *Lagrange multipliers* resulting from the maximization of Eq. (A1) under the chosen constraints coded in the components of the vector $\vec{C}$.

Second, we proceed to calculate the exact values of the Lagrange multipliers. Following Squartini and Garlaschelli [43] and Garlaschelli and Loffredo [45], we determine these parameters by imposing that the properties of the real network are found in the ensemble with maximum probability. Indeed, we may write the log-likelihood of observing the real network $L(\vec{\theta}) = \ln[P(G^*|\vec{\theta})]$ as

$$L(\vec{\theta}) = -H(G^*,\vec{\theta}) - \ln Z(\vec{\theta}). \qquad (A3)$$

Maximizing this quantity thus allows us to fix 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 nonbias [43].

2. Ensemble for a bipartite network with constrained degree sequences

We now explain how the randomizing scheme by Squartini and Garlaschelli [43] applies to our specific problem, namely, a *bipartite network* subject to *local* constraints. The scheme has already been applied to study international trade networks [44].

To begin with, we construct the Hamiltonian for a bipartite network, whose bipartite matrix we call **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 bipartite network, for the sake of clarity, we restrict our notation to the paradigmatic case of plant-pollinator communities. Thus, we speak of systems with N_P number of *plants* and N_A pollinating *animals*. The degree sequences corresponding to the observed network are represented, respectively, by $\vec{v}$ and $\vec{h}$, where v_p is the number of *visiting* animal species that a plant species p receives (degree of the plant species p), while h_a is the number of different *hosting* plant species interacting with the pollinator species a (degree of the animal species a).

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

$$H(\mathbf{B},\vec{\alpha},\vec{\beta}) = \vec{\alpha} \cdot \vec{v} + \vec{\beta} \cdot \vec{h}. \qquad (A4)$$

The introduction of the Lagrange multipliers means that the probability Eq. (A2) of encountering a bipartite graph **B** in the exponential random graph ensemble becomes

$$P(\mathbf{B}|\vec{\alpha},\vec{\beta}) = \frac{e^{-\vec{\alpha}\cdot\vec{v}-\vec{\beta}\cdot\vec{h}}}{\sum_B e^{-\vec{\alpha}\cdot\vec{v}-\vec{\beta}\cdot\vec{h}}}. \qquad (A5)$$

As in Ref. [43], we introduce the variable change $x_p = e^{-\alpha_p}$ and $y_a = e^{-\beta_a}$. Then, the log-likelihood of encountering the real network degree sequences in the ensemble is

$$L(\vec{x}) = \sum_{p=1}^{N_P} v_p \ln(x_p) + \sum_{a=1}^{N_A} h_a \ln(y_a)$$
$$- \sum_{a=1}^{N_A}\sum_{p=1}^{N_P} \ln(1 + x_p y_a), \qquad (A6)$$

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:

$$v_p = \sum_{a=1}^{N_A} \frac{x_p y_a}{1 + x_p y_a} \quad \text{for } p = 1,...,N_P, \qquad (A7)$$

$$h_a = \sum_{p=1}^{N_P} \frac{x_p y_a}{1 + x_p y_a} \quad \text{for } a = 1,...,N_A. \qquad (A8)$$

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).

3. Probability matrix of interactions

It has also been shown in Ref. [43] that in the case of local constraints, the probability of finding a network **B** in the ensemble can be factorized in terms of the probability of the existence of a link between a plant species "p" and animal species "a," which we call p_{pa}. In effect, by taking $p_{pa} = [(x_p y_a)/(1 + x_p y_a)]$, replacing it into Eq. (5), and doing a little algebra, one finds

$$P(\mathbf{B}|\vec{\alpha},\vec{\beta}) = \prod_{p,a} p_{pa}^{b_{pa}}(1 - p_{pa})^{1-b_{pa}}, \qquad (A9)$$

where b_{pa} is the (p,a) element of the bipartite matrix of interactions. Then, using Eq. (A9), it is almost immediate to see that $\langle b_{pa}\rangle = p_{pa}$, thus, in turn, $\langle b_{pa}\rangle = [(x_p y_a)/(1 + x_p y_a)]$. This shows that, as we said, the right-hand side of Eqs. (A7) and (A8) is a sum over a column or row of expected values of the randomized bipartite matrix.

We also note that the possibility of factorizing $P(\mathbf{B}|\vec{\alpha}, \vec{\beta})$ essentially entails that the existence of a link between any pair of species of different guilds constitutes a set of independent events. This factorization is a consequence of the fact that the considered constraints concern average values of local quantities (here the degrees). In this way, one obtains the exact expected randomized matrix of interactions:

$$\langle \mathbf{B}^* \rangle = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1a} & \cdots & p_{1N_A} \\ p_{21} & p_{22} & \cdots & p_{1a} & \cdots & p_{2N_A} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{p1} & p_{p2} & \cdots & p_{pa} & \cdots & p_{pN_A} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{N_P1} & p_{N_P2} & \cdots & p_{N_Pa} & \cdots & p_{N_PN_A} \end{pmatrix}.$$

$$(A10)$$

APPENDIX B: COMPUTATIONAL IMPLEMENTATION

Here we give the numerical details on how we obtain the Lagrange multipliers $\vec{x}^*$ and $\vec{y}^*$ that define the corresponding statistical ensembles of the empirical networks. The determination of these multipliers might be achieved following either of two procedures: by directly maximizing the log-likelihood in Eq. (A6) or by solving the nonlinear coupled set of Eqs. (A7) and (A8). While in this section we present our particular implementation for bipartite graphs, there is a MATLAB package developed by Squartini et al. [47] which numerically solves this optimization problem for a variety of types of unipartite graphs and constraints [69]. The main difference between their implementation and our approach lies in the numerical functions used to find the optimal Lagrange multipliers. Whereas their package uses a local optimizing function [71], we make a special effort to ensure that the maxima we find are global, in particular, by combining the use of a global search algorithm with a local optimization function repeated over a large set of pseudorandom initial conditions.

1. Direct maximization of Eq. (A6)

We numerically optimize the log-likelihood by *simulated annealing* [73–75]. Given the pseudoaleatory character of this approach, which allows us to overcome the barriers separating local minima, it is extendedly used in situations in which the coexistence of several local optima is expected.

More precisely, in our case we need to take into account that in Eq. (A6) 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 degeneracy, we introduce 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 $N'_P = N_P - \mathrm{red}_P$ and $N'_A = N_A - \mathrm{red}_A$. This procedure is an extension to the bipartite case of the redimensionalization proposed in Ref. [45] for a unipartite network. Consequently, the log-likelihood might be rewritten as

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

$$- \sum_{a=1}^{N'_A} \sum_{p=1}^{N'_P} m_p m_a \ln(1 + x_p y_a). \qquad (B1)$$

Although, in analytical terms, the original expression in Eq. (A6) and this latter one are obviously equivalent, from a computational point of view, reducing the number of variables enhances the algorithm's efficiency. In addition, imposing from the beginning such an identity between variables improves the accuracy of the program.

We program a standard version of the simulated annealing algorithm. The random number generator we use is the one by Toral and Chakrabarti [76], with a starting temperature of $T = 10^3$, a reduction factor of the temperature of $\mathrm{RT} = 0.85$, and a total number of updates per fixed temperature of 2×10^4. The algorithm stops when five consecutive iterations differ in less than a parameter $\mathrm{tol} = 10^{-6}$. Furthermore, we run the algorithm ten times per network with different random seeds in order to produce independent sequences of explorations. We consider that the global optimum is reached when all the runs converge to the same solution.

2. Solution of the system given by Eqs. (A7) and (A8)

We solve the set of equations by means of a local deterministic algorithm known as the modified Powell hybrid method. In particular, we use the MINPACK library [77] for FORTRAN available online [78]. This method finds the zero of a nonlinear system by exploiting its Jacobian, which we analytically calculate and implement into the program.

As before, we redimensionalize the problem to N'_P equations for plants and N'_A equations for animals, which now read

$$v_p = \sum_{a=1}^{N'_A} \frac{m_a x_p y_a}{1 + x_p y_a} \quad \text{for } p = 1, ..., N'_P, \qquad (B2)$$

$$h_a = \sum_{p=1}^{N'_P} \frac{m_p x_p y_a}{1 + x_p y_a} \quad \text{for } a = 1, ..., N'_A \qquad (B3)$$

We implement these equations and their Jacobian and run the algorithm with a tolerance $\mathrm{tol} = 10^{-11}$ (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 lies in the risk of getting trapped in local optima, from which, due to its deterministic nature, it is unable to escape. To compensate for this drawback, we perform 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 is not always able to converge to a solution. The rate of success is approximately 50%.

To finally ensure that we find the global maximum, we compare the outcomes of the various independent runs. Moreover, for the cases when the Powell algorithm converges we also compare the solutions obtained for both methods which amounts to a total of ten runs for the simulated annealing and 10^4 for the Powell hybrid method. In all these cases, the same maximum is found.

We also check that the constraints are correctly met with a relative precision between 0.01% and 10% by computing the expected degrees from Eqs. (A7) and (A8) and comparing to the corresponding values of the observed networks. The worst case of 10% is 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. 2 in main text as an example).

APPENDIX C: ANALYTICAL EXPRESSIONS OF THE FIRST TWO MOMENTS OF THE STABLE NODF

The definition of the stable NODF (also named SNODF) is, as proposed by Mariani *et al.* [49], analogous to the classic metric NODF except for the decreasing fill term. In particular, keeping the same notation as in the main text, it reads

$$\text{SNODF}(\mathbf{B}) = \frac{1}{K}\sum_{i<j}^{N_P}\frac{\sum_{a=1}^{N_A} b_{ia}b_{ja}}{v_j} + \frac{1}{K}\sum_{k<l}^{N_A}\frac{\sum_{p=1}^{N_P} b_{pk}b_{pl}}{h_l},$$

(C1)

$$\text{where } K = \frac{N_P(N_P-1)+N_A(N_A-1)}{200}. \quad \text{(C2)}$$

Note that the definition above requires the network $\mathbf{B}$ to be ordered by decreasing degree in both guilds. Having said that, the primary advantage of removing the decreasing fill factor is that the contribution of pairs of rows or columns with exactly the same degree is taken into account. This is specially important for networks where the fraction of repeated degrees is significant either because the number of nodes in one or both guilds is small or because the density of links is high. Nevertheless, for the majority of networks

in our dataset, the values of nestedness measured by NODF and SNODF are closely similar.

The analytical expression for the average of the SNODF over the ensemble is

$$\langle \text{SNODF}(\mathbf{B})\rangle^* = \frac{1}{K}\sum_{i<j}^{N_P}\left\{\frac{\sum_{a=1}^{N_A}\langle b_{ia}\rangle\langle b_{ja}\rangle}{\sum_{a=1}^{N_A}\langle b_{ja}\rangle}\right\}$$
$$+ \frac{1}{K}\sum_{k<l}^{N_A}\left\{\frac{\sum_{p=1}^{N_P}\langle b_{pk}\rangle\langle b_{pl}\rangle}{\sum_{p=1}^{N_P}\langle b_{pl}\rangle}\right\}. \quad \text{(C3)}$$

The standard deviation of the SNODF is given by the analogues of Eqs. (11) and (12), where the partial derivatives in Eq. (12) correspond to

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

(C4)

$$K\frac{\partial\,\text{SNODF}(\mathbf{B})_{\text{animals}}}{\partial b_{rc}} = \sum_{l=c+1}^{N_A}\frac{b_{rl}}{h_l} + \sum_{k=1}^{c-1}\frac{b_{rk}}{h_c} - \sum_{k=1}^{c-1}\sum_{p=1}^{N_P}\frac{b_{pk}b_{pc}}{h_c^2}.$$

(C5)

APPENDIX D: SIGNIFICANCE TESTS

We quantify the significance of the nestedness measures using the *z-score* index, which for a general property x reads $[(x^* - \langle x\rangle)/\sigma_x]$. For us, $\langle x\rangle$ is the average nestedness computed in the ensemble, either analytically or by explicit sampling, and we compare it with the empirical observations x^*. The standard deviation is σ_x. Given that the nestedness values are Gaussian distributed in the random ensemble, the z scores can be directly related to p values. We perform a multiple test correction which allows us to account for the fact that as the number of statistical tests increases, so does the probability of finding rare events [79]. Thus, when considering the multiple comparisons, we can 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 employ the false discovery rate method, in particular, the Benjamini-Hochberg procedure which applies to independent tests [79]. The correction is numerically carried out using the STATSMODEL package in PYTHON [80].

TABLE III. Fraction of networks whose discrepancy between the real and randomized nestedness is less than or equal to 1σ or 2σ for nestedness measures performed with the stable NODF.

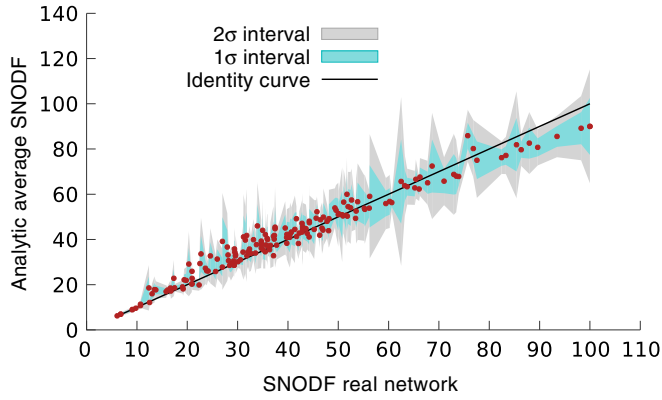| Fraction of ntws with $|z$ score$| \leq 1$		Fraction of ntws with $|z$ score$| \leq 2$	
118 out of 167	70.7%	162 out of 167	97.0%

FIG. 9. Comparison between the real measure of the stable NODF and its average in the ensemble estimated using an analytical expression for the 167 networks of our study.

APPENDIX E: STATISTICAL MEASURES OF THE STABLE NODF

We compute the real nestedness and its statistical significance using the stable NODF for the 167 networks in our dataset. In particular, for each real network we calculate the estimated average and the standard deviation using the analytical expressions in Eq. (C3)–(C5). Figure 9 and Table III show that real nestedness is not statistically significant when measured by the stable NODF.

APPENDIX F: STATISTICAL MEASURES OF THE NODF ON A SAMPLING

We also perform the statistical measures of the NODF on a sampling of the ensemble. The sampling is formed by 10^4 networks generated using the link probabilities in Eq. (A10) for each one of the 167 empirical networks. As we can see in Fig. 10 and Table IV, the real value of nestedness and the average over the sampling are statistically compatible.
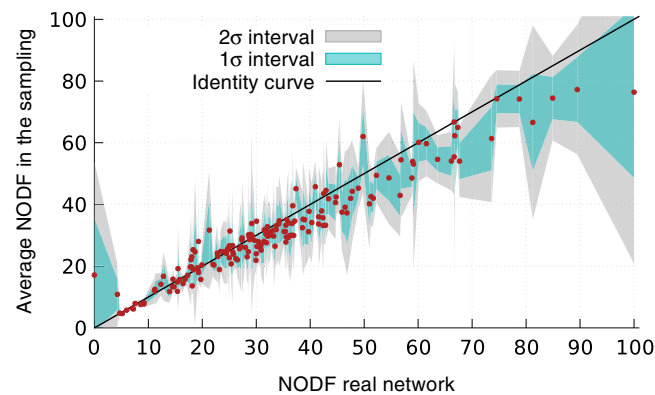

FIG. 10. Comparison between the real measure of the NODF and the average computed over the sampling formed by 10^4 networks for the 167 networks of our study.

TABLE IV. Fraction of networks whose discrepancy between the real and randomized nestedness is less than or equal to 1σ or 2σ for NODF distributions estimated on a sampling of the ensemble.

Fraction of ntws with $|z\text{ score}| \leq 1$		Fraction of ntws with $|z\text{ score}| \leq 2$	
108 out of 167	64.7%	150 out of 167	89.8%

APPENDIX G: STATISTICAL MEASURES OF SPECTRAL RADIUS ON A SAMPLING

We carry out as well the calculations for the expected average and standard deviation of the spectral radius on a sampling. To this end, we sample the statistical ensemble by producing 10^4 networks. For each sampled network, we compute the spectral radius ρ using the R package RARPACK [52]. Finally, we calculate the average and the standard deviation of the resulting distribution. Figure 11 and Table V show the comparison between the real value of the spectral radius and the average computed over the ensemble, which turn out to be statistically compatible.

It is important to remark that the sampling and the theoretical approach provide slightly different measures of nestedness when using the spectral radius. This is due to the non-normalized character of the spectral radius, which sets a dependence on both the size of the network and its number of links. Indeed, a superior bound for the spectral radius which depends on these quantities is given by Yuan [81], who proposed that, for a connected graph A having N nodes and L links, it is fulfilled that $\rho(A) \leq \sqrt{2L - N + 1}$. In fact, when sampling the ensemble, the number of links is conserved on average, but the average number of nodes with nonzero degree decreases. This means that the average density grows and, consequently, so does the spectral radius. This explains the difference between the measures over the sampling and the theoretical ones and calls for
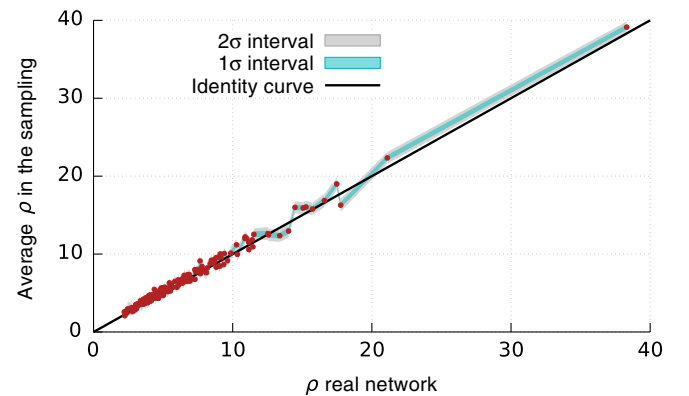

FIG. 11. Comparison between the real observation of the spectral radius $\rho(\lambda)$ and the average estimated over a sampling of the statistical ensemble formed by 10^4 networks for the 167 networks of our study.

TABLE V. Fraction of networks whose discrepancy between the real and randomized nestedness is less than or equal to 1σ or 2σ. Nestedness is measured with the spectral radius $\rho(\lambda)$.

Fraction of ntws with $|z \text{ score}| \leq 1$		Fraction of ntws with $|z \text{ score}| \leq 2$	
84 out of 167	50.3%	149 out of 167	89.2%

special caution when using the spectral radius to quantify nested patterns.

APPENDIX H: SELF-ORGANIZING NETWORK MODEL

In order to reorganize the original network into an even more nested structure, we numerically implement the self-organizing network model proposed by Burgos *et al.* [58]. This methodology keeps constant many aspects susceptible to affecting the measure of nestedness, like the size and fill, but modifies the degree sequences through the redistribution of connections. We rewire the links among species following two simple rules: (i) When changing an interaction, the new partner must have higher degree than the old neighbor, and (ii) if the proposed redistribution leaves one of the two nodes with no interactions at all, we reject the change. This operation is repeated until the system achieves a frozen state in which no more reconnections are accepted (we consider this happens when $10^3 N$ consecutive rejections occur, 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 for this limitation, we carry out 10^3 independent rewiring operations for each network. We then average the target properties, namely, nestedness (measured using the NODF) and the variance of the joint degree sequence of the two guilds.

APPENDIX I: STATISTICAL MEASURES OF DEGREE ASSORTATIVITY

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 use the notion of degree assortativity. We follow the definition proposed by Newman [82], 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 produce for each ensemble a sampling made up by 10^4 networks. We then measure computationally the assortativity of each sampled network using the *assortativity_degree* function from the R IGRAPH package
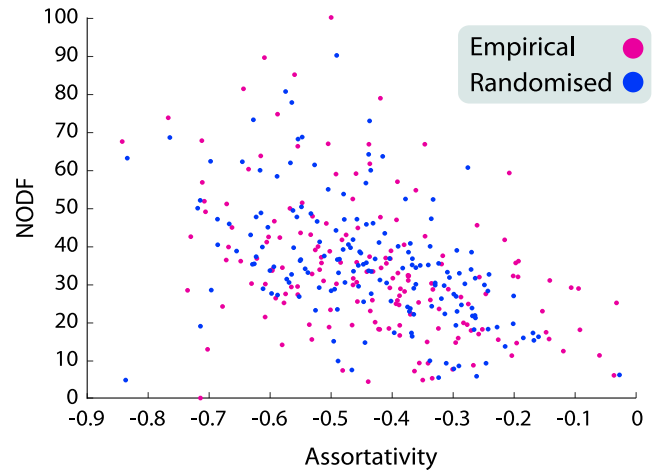

FIG. 12.   Relation between the degree assortativity and the nestedness for the real network (in pink) and the average measure for the randomized case (in blue).

in Ref. [83]. Figure 12 displays the anticorrelation between the Pearson's coefficient and nestedness, showing that, in general, the more dissortativity a network is the more nested it is. Finally, these measures allow us to calculate the first and second moments of the assortativity for each ensemble in our set.

APPENDIX J: DATA

In our study, we analyze 167 real interaction networks from the Web of Life dataset [84]. This set consists of 133 plant-pollinator communities [85–130], 30 seed-dispersal communities [131–148], and four plant-ant communities [149–152]. Data sometime include information about the link's weights, but we convert all networks to binary matrices.

[1] J. Bascompte, P. Jordano, C. J. Melián, and J. M. Olesen, *The Nested Assembly of Plant-Animal Mutualistic Networks*, Proc. Natl. Acad. Sci. U.S.A. **100**, 9383 (2003).
[2] B. D. Patterson and W. Atmar, *Nested Subsets and the Structure of Insular Mammalian Faunas and Archipelagos*, Biological Journal of the Linnean Society **28**, 65 (1986).
[3] S. Saavedra, F. Reed-Tsochas, and B. Uzzi, *A Simple Model of Bipartite Cooperation for Ecological and Organizational Networks*, Nature (London) **457**, 463 (2009).
[4] L. De Benedictis and L. Tajoli, *The World Trade Network*, World Econ. **34**, 1417 (2011).
[5] F. Saracco, R. Di Clemente, A. Gabrielli, and T. Squartini, *Detecting Early Signs of the 2007–2008 Crisis in the World Trade*, Sci. Rep. **6**, 30286 (2016).
[6] L. Hernández, A. Vignes, and S. Saba, *Trust or Robustness? An Ecological Approach to the Study of Auction and Bilateral Markets*, PLoS One **13**, e0196206 (2018).

[7] W. Atmar and B. D. Patterson, *The Measure of Order and Disorder in the Distribution of Species in Fragmented Habitat,* Oecologia **96,** 373 (1993).

[8] R. A. Brualdi and J. G. Sanderson, *Nested Species Subsets, Gaps, and Discrepancy,* Oecologia **119,** 256 (1999).

[9] 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 **117,** 1227 (2008).

[10] J. Galeano, J. M. Pastor, and J. M. Iriondo, *Weighted-Interaction Nestedness Estimator (Wine): A New Estimator to Calculate over Frequency Matrices,* Environ. Model. Softw. **24,** 1342 (2009).

[11] E. Burgos, H. Ceva, L. Hernández, and R. Perazzo, *Understanding and Characterizing Nestedness in Mutualistic Bipartite Networks,* Comput. Phys. Commun. **180,** 532 (2009).

[12] W. Ulrich, M. Almeida-Neto, and N. J. Gotelli, *A Consumer's Guide to Nestedness Analysis,* Oikos **118,** 3 (2009).

[13] A. Solé-Ribalta, C. J. Tessone, M. S. Mariani, and J. Borge-Holthoefer, *Revealing In-Block Nestedness: Detection and Benchmarking,* Phys. Rev. E **97,** 062302 (2018).

[14] D. P. Vázquez and M. A. Aizen, *Community-Wide Patterns of Specialization in Plant-Pollinator Interactions Revealed by Null Models,* in *Plant-Pollinator Interactions: From Specialization to Generalization* (University of Chicago Press Chicago, Illinois, USA, 2006), pp. 200–219.

[15] E. L. Rezende, P. Jordano, and J. Bascompte, *Effects of Phenotypic Complementarity and Phylogeny on the Nested Structure of Mutualistic Networks,* Oikos **116,** 1919 (2007).

[16] 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 (Amsterdam) **394A,** 124 (2014).

[17] M. D. König, C. J. Tessone, and Y. Zenou, *Nestedness in Networks: A Theoretical Model and Some Applications,* Theor. Econ. **9,** 695 (2014).

[18] 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 (London) **458,** 1018 (2009).

[19] E. Thébault and C. Fontaine, *Stability of Ecological Communities and the Architecture of Mutualistic and Trophic Networks,* Science **329,** 853 (2010).

[20] R. P. Rohr, S. Saavedra, and J. Bascompte, *On the Structural Stability of Mutualistic Systems,* Science **345,** 1253497 (2014).

[21] S. Allesina, *The More the Merrier,* Nature (London) **487,** 175 (2012).

[22] S. Allesina and S. Tang, *Stability Criteria for Complex Ecosystems,* Nature (London) **483,** 205 (2012).

[23] J. Grilli, M. Adorisio, S. Suweis, G. Barabás, J. R. Banavar, S. Allesina, and A. Maritan, *Feasibility and Coexistence of Large Ecological Communities,* Nat. Commun. **8,** 14389 (2017).

[24] A. Pascual-García and U. Bastolla, *Mutualism Supports Biodiversity when the Direct Competition is Weak,* Nat. Commun. **8,** 14326 (2017).

[25] C. Gracia-Lázaro, L. Hernández, J. Borge-Holthoefer, and Y. Moreno, *The Joint Influence of Competition and Mutualism on the Biodiversity of Mutualistic Ecosystems,* Sci. Rep. **8,** 9253 (2018).

[26] T. M. Lewinsohn, P. I. Prado, P. Jordano, J. Bascompte, and J. M Olesen, *Structure in Plant-Animal Interaction Assemblages,* Oikos **113,** 174 (2006).

[27] M. Kondoh, S. Kato, and Y. Sakato, *Food Webs Are Built Up with Nested Subwebs,* Ecology **91,** 3123 (2010).

[28] C. J. Melián, J. Bascompte, P. Jordano, and V. Krivan, *Diversity in a Complex Ecological Network with Two Interaction Types,* Oikos **118,** 122 (2009).

[29] S. Kéfi, E. L. Berlow, E. A. Wieters, L. N. Joppa, S. A. Wood, U. Brose, and S. A. Navarrete, *Network Structure beyond Food Webs: Mapping Non-Trophic and Trophic Interactions on Chilean Rocky Shores,* Ecology **96,** 291 (2015).

[30] P. P. Staniczenko, J. C. Kopp, and S. Allesina, *The Ghost of Nestedness in Ecological Networks,* Nat. Commun. **4,** 1391 (2013).

[31] A. James, J. W. Pitchford, and M. J. Plank, *Disentangling Nestedness from Models of Ecological Complexity,* Nature (London) **487,** 227 (2012).

[32] W. Feng and K. Takemoto, *Heterogeneity in Ecological Mutualistic Networks Dominantly Determines Community Stability,* Sci. Rep. **4,** 5912 (2014).

[33] S. Saavedra and D. B. Stouffer, *Disentangling Nestedness Disentangled,* Nature (London) **500,** E1 (2013).

[34] A. James, J. W. Pitchford, and M. J. Plank, *James et al. Reply,* Nature (London) **500,** E2 (2013).

[35] W. Ulrich and N. J. Gotelli, *Null Model Analysis of Species Nestedness Patterns,* Ecology **88,** 1824 (2007).

[36] N. J. Gotelli and W. Ulrich, *Statistical Challenges in Null Model Analysis,* Oikos **121,** 171 (2012).

[37] 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,* J. Theor. Biol. **246,** 510 (2007).

[38] L. N. Joppa, J. M. Montoya, R. Solé, J. Sanderson, and S. L. Pimm, *On Nestedness in Ecological Networks,* Evol. Ecol. Res. **12,** 35 (2010).

[39] S. Jonhson, V. Domínguez-García, and M. A. Muñoz, *Factors Determining Nestedness in Complex Networks,* PLoS One **8,** e74025 (2013).

[40] N. J. Gotelli and G. L. Entsminger, *Swap and Fill Algorithms in Null Model Analysis: Rethinking the Knight's Tour,* Oecologia **129,** 281 (2001).

[41] J. M. Olesen, J. Bascompte, Y. L. Dupont, H. Elberling, C. Rasmussen, and P. Jordano, *Missing and Forbidden Links in Mutualistic Networks,* Proc. R. Soc. B **278,** 725 (2010).

[42] J. Park and M. E. J. Newman, *Statistical Mechanics of Networks,* Phys. Rev. E **70,** 066117 (2004).

[43] T. Squartini and D. Garlaschelli, *Analytical Maximum-Likelihood Method to Detect Patterns in Real Networks,* New J. Phys. **13,** 083001 (2011).

[44] F. Saracco, R. Di Clemente, A. Gabrielli, and T. Squartini, *Randomizing Bipartite Networks: The Case of the World Trade Web,* Sci. Rep. **5,** 10595 (2015).

[45] D. Garlaschelli and M. I. Loffredo, *Maximum Likelihood: Extracting Unbiased Information from Complex Networks*, Phys. Rev. E **78,** 015101(R) (2008).

[46] 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*, Vol. 27 (Kyoto University, 1990), pp. 377–463.

[47] T. Squartini, R. Mastrandrea, and D. Garlaschelli, *Unbiased Sampling of Network Ensembles*, New J. Phys. **17,** 023052 (2015).

[48] J. Borge-Holthoefer, R. A. Baños, C. Gracia-Lázaro, and Y. Moreno, *Emergence of Consensus as a Modular-to-Nested Transition in Communication Dynamics*, Sci. Rep. **7,** 41673 (2017).

[49] M. S. Mariani, Z. M. Ren, J. Bascompte, and C. J. Tessone, *Nestedness in Complex Networks: Observation, Emergence, and Implications*, Phys. Rep. (to be published).

[50] E. Small, *Insect Pollinators of the Mer Bleue Peat Bog of Ottawa*, Can. Field Nat. **90,** 22 (1976).

[51] F. K. Bell, D. Cvetković, P. Rowlinson, and S. K. Simić, *Graphs for which the Least Eigenvalue is Minimal, II*, Linear Algebra Appl. **429,** 2168 (2008).

[52] rARPACK: *Solvers for Large Scale Eigenvalue and SVD Problems*, https://CRAN.R-project.org/package= rARPACK.

[53] E. Deutsch and M. Neumann, *Derivatives of the Perron Root at an Essentially Nonnegative Matrix and the Group Inverse of an m-Matrix*, J. Math. Anal. Appl. **102,** 1 (1984).

[54] S. J. Kirkland and M. Neumann, *Group Inverses of M-Matrices and Their Applications* (Chapman and Hall/CRC, London, 2012).

[55] A. Ben-Israel and T. N. Greville, *Generalized Inverses: Theory and Applications* (Springer Science & Business Media, New York, 2003), Vol. 15.

[56] W. N. Venables and B. D. Ripley, *Modern Applied Statistics with S*, 4th ed. (Springer, New York, 2002).

[57] P. Jordano, J. Bascompte, and J. M. Olesen, *Invariant Properties in Coevolutionary Networks of Plant-Animal Interactions*, Ecol. Lett. **6,** 69 (2003).

[58] E. Burgos, H. Ceva, R. P. Perazzo, M. Devoto, D. Medan, M. Zimmermann, and A. M. Delbue, *Why Nestedness in Mutualistic Networks?*, J. Theor. Biol. **249,** 307 (2007).

[59] S. Johnson, J. J. Torres, J. Marro, and M. A. Munoz, *Entropic Origin of Disassortativity in Complex Networks*, Phys. Rev. Lett. **104,** 108702 (2010).

[60] W. Ulrich and N. J. Gotelli, *A Null Model Algorithm for Presence-Absence Matrices Based on Proportional Resampling*, Ecol. Model. **244,** 20 (2012).

[61] G. Robins, P. Pattison, Y. Kalish, and D. Lusher, *An Introduction to Exponential Random Graph (p*) Models for Social Networks*, Soc. Networks **29,** 173 (2007).

[62] K. Takemoto and M. Arita, *Nested Structure Acquired through Simple Evolutionary Process*, J. Theor. Biol. **264,** 782 (2010).

[63] S. Valverde, J. Piñero, B. Corominas-Murtra, J. Montoya, L. Joppa, and R. Solé, *The Architecture of Mutualistic Networks as an Evolutionary Spandrel*, Nat. Ecol. Evol. **2,** 94 (2018).

[64] D. S. Maynard, C. A. Serván, and S. Allesina, *Network Spandrels Reflect Ecological Assembly*, Ecol. Lett. **21,** 324 (2018).

[65] N. P. Chacoff, J. Resasco, and D. P. Vázquez, *Interaction Frequency, Network Position, and the Temporal Persistence of Interactions in a Plant-Pollinator Network*, Ecology **99,** 21 (2018).

[66] K. Trøjelsgaard, P. Jordano, D. W. Carstensen, and J. M. Olesen, *Geographical Variation in Mutualistic Networks: Similarity, Turnover and Partner Fidelity*, Proc. R. Soc. B **282,** 20142925 (2015).

[67] S. D. Johnson and K. E. Steiner, *Generalization versus Specialization in Plant Pollination Systems*, Trends Ecol. Evol. **15,** 140 (2000).

[68] J. L. Bronstein, R. Alarcón, and M. Geber, *The Evolution of Plant-Insect Mutualisms*, New Phytol. **172,** 412 (2006).

[69] The package implements the entire "Max&Sam" methodology for a general unipartite setting. It provides the functions to find the numerical solution for the maximum entropy ensemble under a number of different constraints for both undirected or directed networks, binary or weighted. Moreover, it permits us to sample the resulting ensemble. The package is freely available online [70]. The bipartite case can be reproduced using this package by redefining the directed unipartite case in such a way that one guild solely has outgoing links, while the other has only incoming links. The formal analogue of preserving the average degree sequences of both guilds is now to constrain, on average, the in-degree and out-degree sequences.

[70] MATLAB Implementation of the Max&Sam Method, https://it.mathworks.com/matlabcentral/fileexchange/ 46912-max-sam-package-zip.

[71] One can find the code in Ref. [70] and the documentation of the MATLAB function they used to solve the optimization problem in Ref. [72].

[72] MATLAB Documentation for the Function "fmincon," https://es.mathworks.com/help/optim/ug/fmincon.html.

[73] A. Corana, M. Marchesi, C. Martini, and S. Ridella, *Minimizing Multimodal Functions of Continuous Variables with the Simulated Annealing Algorithm Corrigenda*, ACM Trans. Math. Softw. **13,** 262 (1987).

[74] W. L. Goffe, G. D. Ferrier, and J. Rogers, *Global Optimization of Statistical Functions with Simulated Annealing*, J. Econom. **60,** 65 (1994).

[75] W. L. Goffe *et al.*, SIMANN: *A Global Optimization Algorithm Using Simulated Annealing*, Studies Nonlin. Dyn. Econom. **1,** 169 (1996).

[76] R. Toral and A. Chakrabarti, *Generation of Gaussian Distributed Random Numbers by Using a Numerical Inversion Method*, Comput. Phys. Commun. **74,** 327 (1993).

[77] J. J. Moré, B. S. Garbow, and K. E. Hillstrom, Technical Report No. CM-P00068642, 1980.

[78] B. S. Garbow, K. E. Hillstrom and J. J. More, MINPACK Library Source Code, Subroutine HYBRD, https://www .math.utah.edu/software/minpack/minpack/hybrd.html.

[79] Y. Benjamini and Y. Hochberg, *Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing*, J. R. Stat. Soc. Ser. B **57,** 289 (1995).

[80] S. Seabold and J. Perktold, STATSMODELS: *Econometric and Statistical Modeling with PYTHON*, in *Proceedings*

of the 9th Python in Science Conference, Vol. 57 (2010), p. 61.

[81] H. Yuan, *A Bound on the Spectral Radius of Graphs*, Linear Algebra Appl. **108,** 135 (1988).

[82] M. E. J. Newman, *Assortative Mixing in Networks*, Phys. Rev. Lett. **89,** 208701 (2002).

[83] G. Csardi, *Assortativity in R IGRAPH Package*, http://igraph .org/r/doc/assortativity.html.

[84] Bascompte Lab, Web of Life, Ecological Networks Database, http://www.web-of-life.es/.

[85] 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*, Am. J. Bot., **69,** 82 (1982).

[86] S. C. Barrett and K. Helenurm, *The Reproductive Biology of Boreal Forest Herbs. I. Breeding Systems and Pollination*, Canadian Journal of Botany (1951-) / Journal Canadien de Botanique **65,** 2036 (1987).

[87] F. E. Clements and F. L. Long, *Experimental Pollination: An Outline of the Ecology of Flowers and Insects* (Carnegie Institution of Washington, Washington, DC, 1923).

[88] L. Dicks, S. Corbet, and R. Pywell, *Compartmentalization in Plant-Insect Flower Visitor Webs*, J. Anim. Ecol. **71,** 32 (2002).

[89] 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 **26,** 301 (2003).

[90] H. Elberling and J. M. Olesen, *The Structure of a High Latitude Plant-Flower Visitor System: The Dominance of Flies*, Ecography **22,** 314 (1999).

[91] J. M. Olesen, L. I. Eskildsen, and S. Venkatasamy, *Invasion of Pollination Networks on Oceanic Islands: Importance of Invader Complexes and Endemic Super Generalists*, Divers. Distrib. **8,** 181 (2002).

[92] J. Ollerton, S. D. Johnson, L. Cranmer, and S. Kellie, *The Pollination Ecology of an Assemblage of Grassland Asclepiads in South Africa*, Ann. Bot. **92,** 807 (2003).

[93] B. Hocking, *Insect-Flower Associations in the High Arctic with Special Reference to Nectar*, Oikos **19,** 359 (1968).

[94] T. Petanidou, *Pollination Ecology in a Phryganic Ecosystem*, Ph. D. thesis, Aristotle University of Thessaloniki, 1991.

[95] J. Herrera, *Pollination Relationships in Southern Spanish Mediterranean Shrublands*, J. Ecol. **76,** 274 (1988).

[96] J. Memmott, *The Structure of a Plant-Pollinator Food Web*, Ecol. Lett. **2,** 276 (1999).

[97] D. W. Inouye and G. H. Pyke, *Pollination Biology in the Snowy Mountains of Australia: Comparisons with Montane Colorado, USA*, Austral Ecol. **13,** 191 (1988).

[98] P. G. Kevan, *High Arctic Insect-Flower Relations the Interrelationships of Arthropods and Flowers atx Lake Hazen, Ellesmere Island, NWT, Canada*, Ph. D. thesis, University of Alberta, 2002.

[99] 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*, Vol. 27 (Kyoto University, 1990), pp. 309–376.

[100] 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*, Arcti. Antarct. Alp. Res. **34,** 233 (2002).

[101] T. Mosquin and J. Martin, *Observations on the Pollination Biology of Plants on Melville Island, NWT, Canada*, Can. Field-Nat. **81,** 201 (1967).

[102] A. Motten, *Pollination Ecology of the Spring Wildflower Community in the Deciduous Forests of Piedmont North Carolina*, Ph. D. thesis, Duke University.

[103] C. McMullen, *Flower-Visiting Insects of the Galapagos Islands*, Pan-Pac. Entomol. **69,** 95 (1993).

[104] R. B. Primack, *Insect Pollination in the New Zealand Mountain Flora*, New Zealand Journal of botany **21,** 317 (1983).

[105] N. Ramirez and Y. Brito, *Pollination Biology in a Palm Swamp Community in the Venezuelan Central Plains*, Botanical Journal of the Linnean Society **110,** 277 (1992).

[106] N. Ramirez, *Biologia de polinizacion en una comunidad arbustiva tropical de la alta guayana venezolana*, Biotropica **21,** 319 (1989).

[107] 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 **59,** 351 (1978).

[108] 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 **147,** 399 (2005).

[109] M. Percival, *Floral Ecology of Coastal Scrub in Southeast Jamaica*, Biotropica **6,** 104 (1974).

[110] A. Montero, *The Ecology of Three Pollination Networks*, Ph. D. thesis, M. Sc. thesis, Aarhus University, 2005.

[111] T. Ingversen, *Plantpollinator Interactions on Jamaica and Dominica: The Centrality, Asymmetry and Modularity of Networks*, Ph. D. thesis, M. Sc. thesis, Aarhus University, 2006.

[112] 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 **29,** 531 (2006).

[113] M. Kato, *Anthophilous Insect Community and Plant-Pollinator Interactions on Amami Islands in the Ryukyu Archipelago, Japan*, Vol. 29 (Kyoto University, 2000), pp. 157–252.

[114] R. Lundgren and J. M. Olesen, *The Dense and Highly Connected World of Greenland's Plants and Their Pollinators*, Arct. Antarct. Alp. Res. **37,** 514 (2005).

[115] M. Bundgaard, *Tidslig og rumlig cariation i et plantebestøvernetværk*, Ph. D. thesis, M. Sc. thesis, Aarhus University, 2003.

[116] Y. L. Dupont and J. M. Olesen, *Ecological Modules and Roles of Species in Heathland Plant-Insect Flower Visitor Networks*, J. Anim. Ecol. **78,** 346 (2009).

[117] S. Bek, *A Pollination Network from a Danish Forest Meadow*, Ph. D. thesis, M. Sc. thesis, Aarhus University, 2006.

[118] L. Stald, A. Valido, and J. Olesen, *Struktur og dynamik i rum og tid at et bestøvningsnetværk på Tenerife, De Kanariske Øer*, Ph. D. thesis, M. Sc. thesis, University of Aarhus, 2003.

[119] D. P. Vázquez, *Interactions among Introduced Ungulates, Plants, and Pollinators: A Field Study in the Temperate Forest of the Southern Andes*, Ph. D. thesis, The University of Tennessee, 2002.

[120] A. C. Montero, *The Ecology of Three Pollination Networks*, Ph. D. thesis, B.Sc. thesis, University of Aarhus, 1998.

[121] K. Yamazaki and M. Kato, *Flowering Phenology and Anthophilous Insect Community in a Grassland Ecosystem at Mt. Yufu, Western Japan*, Vol. 29 (Kyoto University, 2003), pp. 215–318.

[122] T. Kakutani, T. Inoue, M. Kato, and H. Ichihashi, *Insect-Flower Relationship in the Campus of Kyoto University, Kyoto: An Overview of the Flowering Phenology and the Seasonal Pattern of Insect Visits*, Vol. 27 (Kyoto University, 1990), pp. 465–522.

[123] M. Kato and R. Miura, *Flowering Phenology and Anthophilous Insect Community at a Threatened Natural Lowland Marsh at Nakaikemi in Tsuruga, Japan*, Vol. 29 (Kyoto University, 1996), p. 1.

[124] M. Kato, M. Matsumoto, and T. Kato, *Flowering Phenology and Anthophilous Insect Community in the Cool-Temperate Subalpine Forests and Meadows at Mt. Kushigata in the Central Part of Japan*, Vol. 28 (Kyoto University, 1993), pp. 119–172.

[125] I. Bartomeus, M. Vilà, and L. Santamaría, *Contrasting Effects of Invasive Plants in Plant-Pollinator Networks*, Oecologia **155,** 761 (2008).

[126] E. L. Bezerra, I. C. Machado, and M. A. Mello, *Pollination Networks of Oil-Flowers: A Tiny World within the Smallest of all Worlds*, J. Anim. Ecol. **78,** 1096 (2009).

[127] C. N. Kaiser-Bunbury, S. Muff, J. Memmott, C. B. Müller, and A. Caflisch, *The Robustness of Pollination Networks to the Loss of Species and Interactions: A Quantitative Approach Incorporating Pollinator Behaviour*, Ecol. Lett. **13,** 442 (2010).

[128] C. N. Kaiser-Bunbury, D. P. Vázquez, M. Stang, and J. Ghazoul, *Determinants of the Microstructure of Plant-Pollinator Networks*, Ecology **95,** 3314 (2014).

[129] C. Robertson, *Flowers and Insects Lists of Visitors of Four Hundred and Fifty Three Flowers* (Carlinville, USA, 1929).

[130] J. Vizentin-Bugoni, P. K. Maruyama, V. J. Debastiani, L. d. S. Duarte, B. Dalsgaard, and M. Sazima, *Influences of Sampling Effort on Detected Patterns and Structuring Processes of a Neotropical Plant-Hummingbird Network*, J. Anim. Ecol. **85,** 262 (2016).

[131] J. W. Baird, *The Selection and Use of Fruit by Birds in an Eastern Forest*, Wilson Bullet. **92,** 63 (1980).

[132] B. Beehler, *Frugivory and Polygamy in Birds of Paradise*, Auk **100,** 1 (1983).

[133] 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 **134,** 119 (2003).

[134] F. Crome, *The Ecology of Fruit Pigeons in Tropical Northern Queensland*, Wildl. Res. **2,** 155 (1975).

[135] P. Frost, *Fruit-Frugivore Interactions in a South African Coastal Dune Forest*, Acta XVII Congressus Internationalis Ornithologici, Vol. 2 (1980), pp. 71–1184.

[136] M. Galetti and M. A. Pizo, *Fruit Eating by Birds in a Forest Fragment in Southeastern Brazil*, Ararajuba **4.2,** 71 (1996).

[137] B. K. Snow and D. W. Snow, *The Feeding Ecology of Tanagers and Honeycreepers in Trinidad*, The Auk **88,** 291 (1971).

[138] B. K. Snow and D. Snow, *Birds and Berries*, 1988.

[139] A. Hamann and E. Curio, *Interactions among Frugivores and Fleshy Fruit Trees in a Philippine Submontane Rainforest*, Conservation biology : the Journal of the Society for Conservation Biology **13,** 766 (1999).

[140] 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 **32,** 69 (1985).

[141] G. E. Kantak, *Observations on Some Fruit-Eating Birds in Mexico*, Auk **96,** 183 (1979).

[142] F. Lambert, *Fig-Eating by Birds in a Malaysian Lowland Rain Forest*, J. Trop. Ecol. **5,** 401 (1989).

[143] 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*, Am. J. Primatol. **42,** 1 (1997).

[144] A. L. Mack and D. D. Wright, *Notes on Occurrence and Feeding of Birds at Crater Mountain Biological Research Station, Papua New Guinea*, Emu **96,** 89 (1996).

[145] 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 **16,** 173 (1984).

[146] W. Silva, *Patterns of Fruit-Frugivore Interactions in Two Atlantic Forest Bird Communities of South-Eastern Brazil: Implications for Conservation*, in *Seed Dispersal and Frugivory: Ecology, Evolution and Conservation* (CABI Publishing, New York, 2002), pp. 423–435.

[147] N. Noma, *Annual Fluctuations of Sapfruits Production and Synchronization within and Inter Species in a Warm Temperate Forest on Yakushima Island*, Tropics **6,** 441 (1997).

[148] J. Guitián, *Relaciones entre los frutos y los passeriformes en un bosque montano de la cordillera cantábrica occidental* (University of Santiago, Santiago, 1983).

[149] C. R. Fonseca and G. Ganade, *Asymmetries, Compartments and Null Interactions in an Amazonian Ant-Plant Community*, J. Anim. Ecol. **65,** 339 (1996).

[150] D. W. Davidson, R. R. Snelling, and J. T. Longino, *Competition among Ants for Myrmecophytes and the Significance of Plant Trichomes*, Biotropica **21,** 64 (1989).

[151] D. Davidson and B. Fisher, *Symbiosis of Ants with Cecropia as a Function of Light Regime*, in *Ant-Plant Interactions*, edited by C. R. Huxley and D. F. Cutler (Oxford University Press, Oxford, 1991), pp. 289–309.

[152] 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 **106,** 344 (2004).