ARTICLE TYPE

A comprehensive benchmark for structural assessment in bipartite networks

María J. Palazzi[1] | Aniello Lampo[2] | Albert Solé-Ribalta[1,3] | Javier Borge-Holthoefer*[1]

[1]Internet Interdisciplinary Institute (IN3), Universitat Oberta de Catalunya, Barcelona, Catalonia, Spain

[2]Universitat Rovira i Virgili, , Tarragona, Catalonia, Spain

[3]URPP Social Networks, University of Zurich, Zurich, Switzerland

Correspondence
*Corresponding author. Email: jborgeh@uoc.edu

Abstract

The use of null models has been a cornerstone to assess the emergence of many network properties at different levels of organization (micro-, meso- and macroscale). Notwithstanding, the debate around which is the most appropriate randomization procedure for a given problem is far from being over. Within the ecological community, for example, the discussion around whether nestedness is –or is not– a frequent pattern in natural systems, and under which assumptions, remains open. For this particular problem, efforts have been devoted to exploring to what extent current models are vulnerable to statistical errors, or to introduce new models that employ different randomization procedures. However, few or no attention has been devoted to the performance of those null models against other architectures. Here, we show that assessing alternative structures under a single null model may produce ambiguous results, which difficult the comparison regarding the joint emergence of different arrangements within a single network. To this aim, we analyze the statistical significance –in terms of z-scores– of nestedness, modularity, and in-block nestedness scores, employing five different null models on a benchmark of $\sim 2.5 \times 10^4$ synthetic bipartite networks with prescribed levels of the mentioned patterns. We show that some null models systematically over- or underestimate the presence of one or another structural pattern. In light of these ambiguities, we introduce an alternative model (termed Corrected Probabilistic model) that reduces the observed biases towards under- and overestimation, and highlight the need for the development of new frameworks that take into account those biases.

KEYWORDS:
Bipartite networks; null models; nestedness; modularity; in-block nestedness

1 | INTRODUCTION

Almost twenty years ago, research in community ecology was deeply challenged by the evidence that biota interactions, rather than being randomly assembled, manifest clear architectural patterns. In 2003, the pioneer paper of Bascompte *et al.*[1] showed that a significant amount of mutualistic plant-animal networks are organized according to well-defined nested arrangements. Along the same line, Olesen *et al.*[2] unveiled the prevalence of modular configuration among pollination communities with a high (> 150) number of species. In the following years, both nestedness and modularity were recognized to strongly affect some beneficial properties –the former was linked to the promotion of diversity[3] and the maximization of abundances[4], while

the latter was associated to the enhancing of stability [5,6]– leading to a great consensus about the crucial role they play in the dynamics of natural systems [7,8,9,10,11,12,13,14,15].

The importance of interaction structure in the behavior of ecological communities fueled, consequently, a long series of studies aimed to assess the statistical significance of the emergent patterns. Herein, null models stand up as one of the ubiquitous tools to address this class of problems, in particular referring to nestedness [16,17]. Paraphrasing Gotelli and Graves [18], a null model is a pattern-generating model, designed with respect to some ecological or evolutionary process of interest, which allows for randomization test of ecological data, for instance, a species' network (as in our case). Certain elements of the network are held constant (constraints), and others are permitted to vary stochastically to create new assemblages. In this way, a null model generates a set of objects that match the original one only in the selected constraints and may be employed as a term of comparison to infer wether the network in question displays nontrivial features, such as the architectural pattern, as a consequence of those constraints, namely as a result of a null hypothesis embodying a particular ecological mechanism.

The selection of these constraints has important effects, and yields many controversial aspects concerning the use of null models. On the one hand, dealing with too loose constraints could misguide us to think that a structural pattern really exists (type I error) [19,20]. On the other hand, imposing too strict constraints will increase the chance of falsely rejecting the statistical significance of the considered pattern (type II error) [16,19,21]. For instance, if the null hypothesis refers to species interactions, the first point of view could translate into an ensemble of networks in which degrees are proportionally constrained, matching only on average the marginal total of the original matrix [1]. The second approach, instead, might fix node degrees by swapping pairs of interactions such that both row and column totals are preserved [22]. Intuitively, the former algorithms avail themselves of ample room to generate random structures, thus often inducing a type I error. While the latter will tend to generate null matrices that closely resemble the original network and so, in order to get a truly representative ensemble, a very large number of null matrices needs to be produced, something which is not possible in some scenarios and could lead to the generation of a biased sample.

Since the 70s, scholars have struggled to overcome the aforementioned limitations [23,24]. Later, a myriad of papers have addressed the ambiguities in pattern assessment, both improving existing algorithms, and introducing new ones [25,26,27,28,29]. Many of these, however, present some limitations that need to be further explored or extended, both on the conceptual and technical side. First, many of the works about null models take into account a reduced number of real world networks [30,31,32]. Such procedure implies, to some extent, circularity. For example, testing for nestedness via randomization on an empirical network for which we would like to know whether it presents nestedness or not. Some authors have considered large synthetic networks ensembles, *e.g.*, Strona *et al.* [21] employed an ensemble of 10^3 synthetic networks, but it only considered square networks with up to 50 rows and columns. In general, there is still a lack of a proper synthetic benchmark aimed to build prescribed architectural patterns. The assessment of nestedness (or any other pattern, for that matter), indeed, is often performed over random matrices, and the few papers dealing with effective nested matrices generate them in a contrived way. One exception, at the best of our knowledge, is provided by the work of Beckett *et al.* [33].

Moreover, another pressing issue relates to the structural arrangements under consideration. So far, the debate has focused almost exclusively on nestedness, while modularity, for example, has been mostly left aside with few exceptions [34,35]. Considering the ecological implications of exhibiting one or another pattern, null models should be balanced enough to discriminate correctly among the different possibilities. Furthermore, note that, in addition to nestedness and modularity, there are other architectural patterns worth being considered, *e.g.* in-block nestedness, which describes networks compound by weakly interlinked blocks with internal nested character [36,37,38]. This has recently attracted a lot of attention to be predominant in real mutualistic communities [39] and to emerge as a consequence of micro- [40,41] and macroscopic [42] mechanisms.

In the current manuscript, we first introduce a network generation benchmark which enables the construction of synthetic graphs with varying levels of specified planted structures (from nested, to modular, to in-block nested, to completely random patterns). On top of such benchmark, we analyze the statistical significance –in terms of z-scores– of three rich sets of bipartite matrices with nested, modular and in-block nested configuration. We consider five different null models and detect the presence of ambiguities in assessing the various network descriptors, corresponding to over- (or under-) estimations of the given arrangements. To overcome the observed inconsistencies, we introduce an alternative approach, the Corrected Probabilistic model, that helps to reduce the observed under- and overestimation, and highlights the need for the development of new frameworks that take into account those biases. These contributions are wrapped into a single compact framework, consisting of software repositories that enable the generation of the synthetic networks with the specified planted structures, the calculation of nestedness under different formulations, optimization algorithms for the maximization of modularity and in-block nestedness, and the construction of the reviewed null models.

The article is organized as follows: Section 2.1 provides the definitions for nestedness, modularity, and in-block nestedness. Section 2.2 provides the details of the suitable synthetic network generation benchmark, that is employed in this work. Section 2.3 offers details regarding the empirical data we analyzed. Section 2.4 introduces the five null models that will be studied in this work. Section 3 presents the results of our study in both the synthetic and empirical benchmarks. Finally, the discussion in Section 4 summarizes all the previous and highlights the main open challenges for further developments.

2 | MATERIALS AND METHODS

2.1 | Relevant structural patterns in Ecological Networks

We present here the three structural arrangements considered in the manuscript: nestedness at the macroscale, modularity and in-block nestedness at the mesoscale. We provide their definition and discuss the related properties, paying particular attention to the measures aimed to quantify them –implemented and available at https://github.com/COSIN3-UOC/N-Q-IBN.
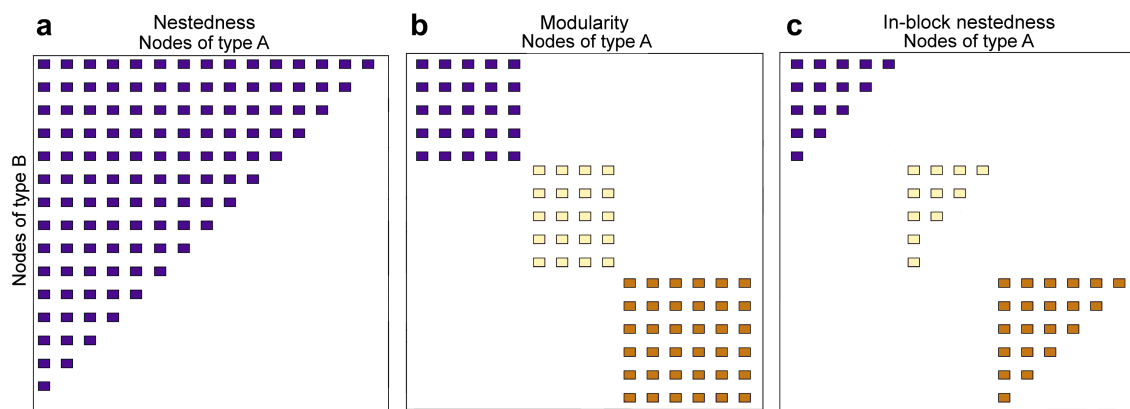

FIGURE 1 Idealised examples of adjacency matrices associate to nested (**a**), modular (**b**), in-block nestedness (**c**) networks. For perfect nested networks, the adjacency matrix manifests a triangular structure, while in the modular case it is divided in blocks, exhibiting a high internal connectivity and a low one. The adjacency matrix of an in-block nested matrix is divided in blocks with an internal nested structure.

2.1.1 | Nestedness

Nestedness, originally conceived in the field of biogeography[43,44,45], constitutes a frequent architectural pattern both inside[1,13] and outside Ecology[46,47,48,49,15]. A perfect nested configuration describes a hierarchical organization where the set of neighbors of a node is a subset of the neighborhoods of larger degree nodes. That is, given a graph G and let V be the set of its vertexes, a perfectly nested configuration exists if

$$\forall i, j \in V, \quad k_i < k_j \iff \Gamma_i \subset \Gamma_j, \tag{1}$$

where k_n and Γ_n indicate the degree of a node n and its neighborhood, respectively[1]. In terms of adjacency matrixes, this definition translates into the characteristic triangular shape, as that shown in Fig. 1 a.

Even though perfectly nested patterns are extremely rare in real complex networks, many of these exhibit a highly nested character. Quantifying nestedness has rendered itself a large amount of definitions[50]. From an algebraic perspective, the spectral properties of perfectly nested graphs have been studied by mathematicians[51,52] facilitating the proposal of the detection method of Staniczenko *et al.*[11], which quantifies nestedness with respect to the maximum eigenvalue of binary and weighted graphs' adjacency matrices. Ecologists have also dedicated many efforts to evaluate nestedness in real systems. In first place,

---

[1]In the case of bipartite networks, this definition is valid provided one compares nodes belonging to the same class.

there are measures based on counting misplaced relations to complete a perfect upper triangular nested structure in the biadjacency matrix, such as the Nested Temperature (NT) measure, introduced by Atmar and Patterson[53]. To overcome some pitfalls around placement-based measures, Almeida-Neto et al.[54] developed overlap metrics, like the Node Overlap and Decreasing Fill (NODF), which considers the amount of common neighbours between every two pairs of nodes in matrix A, alongside with its weighted version[55,56,57].

In this work, we stick to a NODF-like descriptor, which for the case of a bipartite network writes as

$$\mathcal{N} = \frac{2}{N_r + N_c} \left\{ \sum_{i,j}^{N_r} \left[\frac{O_{i,j} - \langle O_{i,j} \rangle}{k_j(N_r - 1)} \Theta(k_i - k_j) \right] + \sum_{l,m}^{N_c} \left[\frac{O_{l,m} - \langle O_{l,m} \rangle}{k_m(N_c - 1)} \Theta(k_l - k_m) \right] \right\}, \tag{2}$$

where N_r and N_c represent the number of adjacency matrix rows and columns, respectively; $O_{ij} = \sum_k a_{ik} a_{jk}$, and similarly O_{lm}, account for the amount of commonly shared neighbours –a.k.a overlap– between pair of species (row or column node pairs); $k_i = \sum_k a_{ik}$ corresponds to the degree of node i and quantifies the number of species with whom i is related to; and $\Theta(\cdot)$ is the Heaviside step function, that ensures that O_{ij} has a positive contribution when $k_i \geq k_j$. Additionally, O_{ij} is conveniently corrected by a null model that discounts the expected change of each species have to share a neighbor[38], namely, the expected overlap $\langle O_{ij} \rangle$. Assuming no correlation between neighboring species of i and j the probability of sharing a particular neighbour only depends on the degree of i and j and on size of the network, $(k_i k_j)/N_c^2$. Hence, the average overlap is $\langle O_{ij} \rangle = \sum_{k=1}^{N_c} (k_i k_j)/N_c^2 = (k_i k_j)/N_c$. The same argument applies for the other set of species $(k_l k_m)/N_r$. Note that Eq. 2 follows closely (and often reduces exactly) to NODF (Almeida-Neto *et al.*) if the term $\langle O_{ij} \rangle$ is suppressed. We chose to employ this metric to assess nestedness, instead of the original NODF since it also incorporates a term that discounts the expected overlap among pairs of nodes. Notwithstanding, as we will show in Sec.3, our results for the statistical significance of nested networks over the different null models are robust against other existing metrics, *i.e.* spectral radius, and the original NODF.

2.1.2 | Modularity

Modular networks are those built up from weakly interlinked groups (also called communities, modules, compartments, clusters) with high internal connectivity. Fig. 1 b depicts the adjacency matrix of a perfectly modular network, *i.e.* with no intergroup links. Modular structure is a rather ubiquitous mesoscale architecture[58,59,60,61] and appears in several fields, as diverse as biology[59] and cognitive science[62].

The problem of identifying community structure constitutes itself a sub-area of network science. Over the last decades, scholars have developed a rich collection of algorithms and methodologies to infer these communities from relational data[61,63]. Probably, the most popular method in network science, and in particular in Ecology, is through the maximization of a fitness function called modularity $Q \in [0, 1)$[64]. We will also employ this function in the present work. Formally, the modularity measure can be written as

$$Q = \text{argmax}_{\vec{\alpha}} \frac{1}{L} \sum_{i=1}^{N_r} \sum_{j=1}^{N_c} \left(A_{ij} - \frac{k_i k_j}{L} \right) \delta(\alpha_i, \alpha_j) \tag{3}$$

where A is the biadjacency matrix, L is the total number of links in the network, $\vec{\alpha}$ is a vector representing the membership variable of species with entries α_i, $\delta(\alpha_i, \alpha_j)$ corresponds to the Kronecker delta, equal to one if nodes i and j belong to the same community (zero otherwise) and $k_i k_j/L$ is a regularisation term to discount the expectation that two nodes are connected by chance considering the degree they have.

2.1.3 | In-block nestedness

Over the last decade, nestedness and modularity have concentrated most of the attention in community ecology, as they seem to have deep implications on the dynamics and other ecological properties. Nestedness, indeed, promotes the diversity of species[3], but restrains instead the local asymptotic stability[9,11], a quality that is maximized by modular organization[10,6]. This suggests that natural systems, which are often observed to be diverse and stable at the same time, develop jointly nestedness and modularity.

The existence of concurrent nested and modular organizations has been debated in different situations[2,49], but empirical[34] and later analytical[65] evidence proved that macroscale nestedness and modular mesoscale cannot coexist easily[65]. This apparent incompatibility is due to the fact that the two patterns arise as a consequence of different mechanisms: certain pressures promote the block organization, while others favor the emergence of nested patterns. Actually, the analysis of these mechanisms suggests

that the co-occurrence of nestedness and modularity may take place at different scales. Herein, in-block nestedness stands out as a hybrid architecture that results from recasting nestedness at the mesoscale, *i.e.*, within modules [37,38].

In-block nestedness depicts a network configuration where weakly connected blocks exhibit an internal nested assembly (see Fig. 1 c), and was first introduced by Lewinson *et al.* [36]. This has been found to unfold from an abundance-maximization process [4] on top of a niche-structured population [40,41], providing a bottom-up perspective on nestedness-modularity, able to trade off stability and diversity in mutualistic communities [39]. Significantly, in-block nestedness has been shown to be the predominant pattern in a significant amount of real plant-animal communities [39], as well as outside of ecology [66].

The original approaches [37,67,36] to detect nested compartments started first optimizing modularity, then subsequently computing the level of nestedness exclusively for nodes within the detected communities. Although this sequential procedure delivers good results in many situations (since, often, detected modules gather nodes with degree heterogeneity), a specialized fitness function is required for the general case. Thus, as in the case of modularity, in-block nestedness detection is a hard computational problem where the use of heuristic algorithms is mandatory. Using the formulation developed in [38], the degree of in-block nestedness of a network $\mathcal{I}$ can be computed as

$$\mathcal{I} = \frac{2}{N_r + N_c} \left\{ \sum_{i,j}^{N_r} \left[\frac{O_{i,j} - \langle O_{i,j} \rangle}{k_j (C_i - 1)} \Theta(k_i - k_j) \delta(\alpha_i, \alpha_j) \right] + \sum_{l,m}^{N_c} \left[\frac{O_{l,m} - \langle O_{l,m} \rangle}{k_m (C_l - 1)} \Theta(k_l - k_m) \delta(\alpha_l, \alpha_m) \right] \right\}, \quad (4)$$

where C_i and C_l are the size of the communities where nodes i and l belongs to. Worth highlighting this hybrid structure reframes nestedness, originally a macroscale feature, to the mesoscopic level. In this sense, a perfectly nested structure corresponds to an in-block nested structure with a single community, i.e. Eq. 4 reduces *exactly* to Eq. 2 when $B = 1$.

2.2 | Synthetic benchmark

We carry out the analysis of the robustness of different types of null models over three sets of synthetic networks with the three patterns of interest. So far, many of the works concerning the study of structured interactions in Ecology proceed by quantifying the degree of a given pattern over networks with a random organization, or nested networks constructed in contrived way. The only exception, to the best of our knowledge, is constituted by Beckett and Williams [33], who construct synthetic perfect nested matrices with varying levels of connectance, by employing the unit ball equation, and then iteratively perform rewires to degrade the ideal structure, introducing noise. Thus, one of the main contributions of the current manuscript is the generation of networks with a well-defined structure, as a variation of the model introduced in [38], and available at https://github.com/COSIN3-UOC/network-generator. In our framework, the different architectural configurations are controlled by four tunable parameters: the number of communities $B \in [1, \infty]$, the inter-community noise $\mu \in [0, 1]$, the noise outside the perfect nested structure $p \in [0, 1]$, and the shape parameter controlling the slimness of the nested profile $\xi \in [1, \infty]$.

The construction of an $N \times N$ adjacency matrix (the extension to rectangular matrixes is straightforward) with a given structure starts off by fixing the number of blocks. Particularly, we build $\lfloor B \rfloor$ blocks of size $\lfloor N/B \rfloor$ and another with the remaining $N - \lfloor N/B \rfloor$ nodes, where $\lfloor . \rfloor$ represents the integer part function. The intra-block structure is obtained by first generating a perfect nested configuration, which shape depends on the parameter ξ as

$$y = f_n(x) = 1 - (1 - x^{1/\xi})^\xi, \quad x \in [0, 1], \quad (5)$$

inspired by the unit ball equation. Specifically, a block with N_B nodes is constructed by tessellating the $[0, 1] \times [0, 1]$ space into N_B^2 squares and adding a link into each entry whose center lies above the curve in Eq. 5. The resulting perfect nested blocks are rarely found in real systems, thus we introduce a noise relying on a dual-step procedure: a fraction p of links lying above the curve in Eq 5 are randomly selected and then distributed across the empty entries of the block. Similarly, we reproduce inter-block perturbations with an additional noise procedure ruled by the parameter μ: for each block, each link is removed with probability $p_i = \mu(B - 1)B$ and then distributed at random to connect a node of the original block to another of a different one.

Figure 2 shows several synthetic networks related to different values of the model parameters. The first row refers to nested matrices, which can be obtained by setting $B = 1$ and $\mu = 0$, *i.e.* considering only one block, and are fully characterized by the choice of ξ and p. From left to right, the former parameter takes increasing values and the resulting nested shape becomes more and more slight. The value of p, instead, decreases from left to right. The situation in which $p = 1$ yields to a Erdös-Rényi graph. The middle row presents a few examples of modular matrices, exhibiting a different number of blocks B. These are obtained by setting $p = 1$ in order to get rid of any intra-block structure, while ξ is fixed according the wished edge density. The parameter μ allows to interpolate between a purely modular network ($\mu = 0$), *i.e.* without inter-blocks links, and an Erdös-Rény one ($\mu = 1$).

Finally, the third row portrays in-block nested matrices associated to several parameter choices, paying particular attention to the role of p and μ. In the limit in which $p = \mu = 0$ we have a purely in-block nested configuration, while for $p = \mu = 1$ an Erdös-Rény arises.

The generative process described above may be implemented in terms of edge probabilities, simplifying the process of algorithmic design. The probability of having a link between nodes i and j within a block writes as

$$P_{ij}^{(intra)} = \left[(1 - p + pp_r)\Theta(jN - f_n(iN)) + p_r(1 - \Theta(jN - f_n(iN))) \right] (1 - p_i). \tag{6}$$

The first term within square brackets, $(1 - p)$, embodies the probability of not altering a link, while the second, pp_r, provides that of recovering a link after removal in the random dispersion of removed links. The Heaviside function restricts them to the noiseless block portion. The third term, $p_r = pE(N - E + pE)$ (E edge number) gives the probability of selecting links in the random distribution of removed ones. The probability of an inter-block link is

$$P_{ij}^{(inter)} = \frac{2Ep_i}{2(B-1)N^2} = \frac{\mu E}{N^2 B}, \tag{7}$$

where the numerator takes into account the amount of removed links, while the denominator regards the possible matrix entries where those links can be placed. For further details regarding the network generation model we refer the reader to [38].
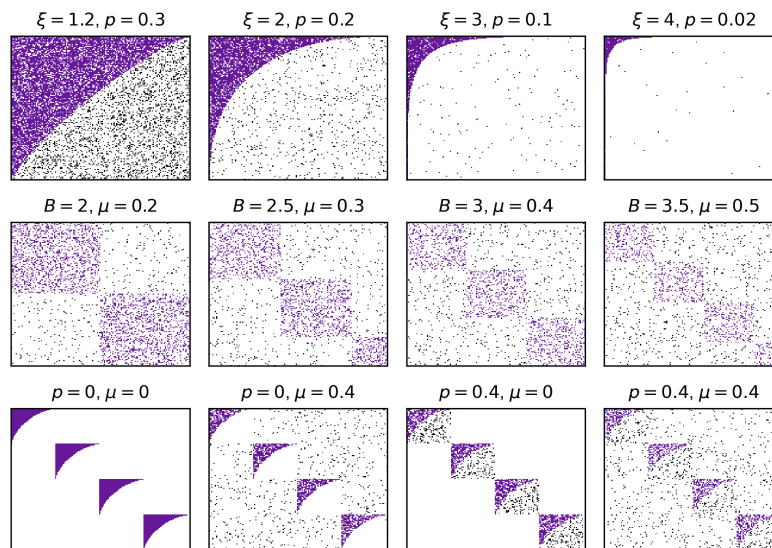

FIGURE 2 Examples of synthetic network generation with the model introduced in [38]. The top row shows the results for the generation of nested networks ($B = 1$) and the effects of varying ξ and p parameters. Middle row exhibits results for modular networks with varying levels on the number of blocks B and the inter-block parameter μ. Finally, the bottom row corresponds to in-block nested networks, and provides some examples of the effect of the noise parameters p and μ.

2.3 | Empirical datasets

To explore our hypothesis in real scenarios, we perform experiments on 192 bipartite ecological networks from two papers and two online databases: 35 trophic networks from [68,69], 141 mutualistic networks from the Web of Life dataset [70], and 16 (3 mutualistic and 13 trophic) networks from the Interaction Web Database http://www.ecologia.ib.usp.br/iwdb hosted by the National Center for Ecological Analysis and Synthesis (University of California). All the networks in our dataset were treated as binary (unweighted links) and we have only kept in our study networks with a minimum of 20 nodes, to ensure that the identified architectural patterns are well defined. In parallel, to reduce computational costs we restricted our analysis to networks below

300 nodes. We highlight, however, that the number of nodes in most of the networks lie between this range, values above or below can be considered outliers, see Section 3.3 below.

2.4 | Null Models

The appropriate selection of a null model has been a center of debate in the ecology literature for many decades. The first attempts to estimate the significance of structural patterns in ecological networks date back to the 70s[23,24]. In this context, scholars introduced approaches relying on the generation of random networks, either by keeping fixed the number of rows (or columns) nodes degrees and allowing the number of columns (or rows) nodes degrees to change[23]; or by generating networks that exactly preserved degrees of both guilds[24]. Afterward, models that preserves, on average, nodes degrees in both guilds were also introduced[71,1]. More recently, models based on constructing maximum-entropy ensembles of networks with an expected degree sequence were also exploited to assess statistical significance[48,72].

In Gotelli *et al.*[22], the authors considered three basic constraints that can be imposed independently for rows or columns nodes: 1) where the individual nodes' degrees are exactly preserved, namely *fixed* (F), 2) where the nodes' degrees are preserved on average, i.e., the probability of interaction is *proportional* (P) to the degree, and 3) where all the pairwise interactions between nodes are *equally probable* (E). Resulting in a total of $3^2 = 9$ possible null models, each one having multiple possible implementations.

Below, we provide details for these three basic classes of null models and describe some of the implementations for each case, with special attention to those that were employed in the present work. It worth highlighting that in all of our implementations, the corresponding constraints were applied for both rows and columns nodes. Therefore, we will talk about EE, FF, and PP models, according the nomenclature employed in[16].

2.4.1 | Equiprobable-Column or Equiprobable-Row (EE) model

The least constrained scheme reviewed here, EE is a probabilistic binary null model that assumes that all the interactions between pairs of nodes –row-to-column links– are equally probable, and that such connection is given with probability $p_{ij} = \frac{L}{(N_r N_c)}$, where L is the total number of connections in the original matrix. This model generates Erdös-Rényi networks with nodes exhibiting a varying number of connections around a well-defined average degree $\langle k \rangle$, preserving L on average. It is known for its tendency to falsely detect significant nested patterns, *i.e.*, it is prone to type I statistical errors[19,73,20].

2.4.2 | Fixed-Column Fixed-Row (FF) models

In the other extreme of the spectrum[74], the FF model enforces row-nodes' and column-nodes' degrees to be strictly preserved. It is thus the most restrictive null model. The usage of this type of model to assess the probability that an observed pattern could be explained by chance alone has been applied since the 70s[24]. A typical way for generating random matrices under this constraint is through the application of some form of swap algorithm, performing exchanges between pairs of interacting nodes without altering the network's degree sequence[22].

Noteworthy, generating random matrices by exactly preserving the degrees is far from being a trivial task. Swap methods tend to generate repeated matrices that closely resemble the original network. Thus, in order to produce a truly representative set of matrices within the universe of possible matrix configurations, a very large number of matrices need be generated[75], which might not be possible in some scenarios. As a consequence, this model is susceptible to type II errors, particularly, when it comes to the assessment of nested patterns[16,73]. On the other hand, it also tends to generate matrices with more checkerboard units[2] than expected by chance[25], producing a biased sample. To solve the latter downside, scholars have introduced variations of the swap methods, and alternative methods that are able to generate a uniform sample of null matrices with fixed row and column degrees[25,26,27,28,29].

In this work, we employ the Curveball algorithm, introduced in[26] as an FF model, as a fast and efficient alternative to generate random matrices with fixed nodes' degrees guaranteeing a uniform sample[76]. The algorithm was introduced as a computationally cheap method to perform permutations on binary matrices, producing uniformly distributed random matrices with fixed degrees. The functioning of the algorithm goes as follows: starting from a binary bipartite matrix –*e.g.*, representing species per site or plant-pollinator matrices, in the ecological context–, for each row (column) node, a list is created containing all the column

---

[2]2×2 submatrices containing 1's in the main diagonal and 0's outside, or vice versa

(row) indices with nonzero elements. Then, the algorithm iteratively chooses at random two row lists (columns), identifies the column (row) indices that are present in list, but not in the other, and vice-versa, and performs random "trades" of the column (row) indices. The number of trades for each pair of lists will go from zero to the size of the smallest list of column indices. In this work, we resort to the Python implementation of the algorithm that was provided by the authors in [26].

2.4.3 | Row-Proportional Column-Proportional (PP) models

Somewhere between the extremes –and limitations– described above[74], an intermediate alternative is to generate models that preserve degrees on average, mitigating then, the proneness to type II error that is present in the FF models, and the proneness to type I error present in the EE model. For this type of model, the probability of interaction of a pair of nodes is assigned to be proportional to the nodes' degrees. From this general principle, there exist different flavors along the loose-strict continuum. Below we introduce three variants employed here.

Bascompte's probabilistic model

One of the most popular probabilistic models for the assessment of statistically significant structural patterns (particularly nested ones) was introduced by Bascompte *et al.* in 2003[1]. In this model, the probability of an interaction by a pair of nodes is proportional to the nodes' degrees of the original matrix as:

$$p_{ij}^{(b)} = \frac{1}{2} \left(\frac{k_i}{N_c} + \frac{k_j}{N_r} \right), \tag{8}$$

where k_i is the degree of the $i-$th row, k_j is the degree of the $j-$th column and N_r and N_c are the respective number of rows and columns nodes.

Maximum-entropy exponential random graphs

Introduced in the last decade[77], this is a statistical physics framework that facilitates the analytic computation of the expected properties of an ensemble of networks. Under this formulation, the expected values and standard deviation of the network properties are calculated by finding a probability distribution over an ensemble of networks –which, on average preserve the network's degree sequence–, that maximizes the entropy of the network. This approach can be applied to both uni- and bipartite networks, and has been recently exploited to assess the significance of nestedness structural arrangements in economic[78], and ecological[72] networks.

With this approach, one aims to find a probability distribution $P(G)$ over the ensemble of graphs G, with G^* as the graph corresponding to the real network, that maximizes the Shannon-Gibbs entropy $S = -\sum_G P(G) \ln P(G)$, and that keeps the average nodes' degree fixed. This maximization has a solution given by the canonical distribution

$$P(G) = \frac{e^{-H(G,\theta)}}{Z(\theta)}, \tag{9}$$

where Z is the partition function, $H(G, \theta) = \theta \cdot C(G)$ is the graph Hamiltonian, and θ is a vector of Lagrange multipliers resulting from the maximization of the Shannon-Gibbs entropy, under the chosen constraints C, e.g., the graph's average degree. The next step consists on the calculation of the exact values of the Lagrange multipliers. Following the approach presented in[77], these multipliers are determined by imposing that the chosen constraints of the network are found in the ensemble with maximum probability. This is achieved by rewriting the log-likelihood of observing the real network as $L(\theta) = -H(G^*, \theta) - \ln Z(\theta)$ and maximizing this quantity in order to find the optimal variables θ^* that define the ensemble. Once the parameters θ^* are found, it is possible to build the matrix containing the average probability of interaction corresponding to our empirical network $\langle G^* \rangle$. Finally, it is possible to derive an analytical expression of the first and second moments of the desired network property. In this work, instead of analytically deriving such statistical properties, we will generate a random ensemble directly from the matrix of average probabilities of interactions.

Corrected Probabilistic Model

Lastly, we now introduce an alternative PP variant that aims to mitigate the possible weaknesses of the models described above. The corrected probabilistic model, in the same spirit as Bascompte's one, sets the probability of interaction by a pair of nodes to be proportional to the nodes' degrees of the observed matrix, but includes two corrections terms. These terms help to reduce the discrepancies observed in the probabilistic model at preserving the nodes' degrees –on average–, while avoiding the production of network instances that are too similar to the original matrix, as in the Maximum-entropy model. In this work, we will test in

depth the robustness of this model against type I and type II errors. Overall, in this model, the threshold ruling the existence of a link is given by:

$$p_{ij}^{(C)} = p_{ij}^{(b)} + \frac{1}{2}\left[\left(1 - \frac{k_i}{N_c}\right)\left(\frac{k_j - \langle k^r \rangle}{N_r}\right)\right] + \frac{1}{2}\left[\left(1 - \frac{k_j}{N_r}\right)\left(\frac{k_i - \langle k^c \rangle}{N_c}\right)\right] \tag{10}$$

where k_i is the degree of the i−th row, k_j is the degree of the j−th column, $\langle k^r \rangle$ and $\langle k^c \rangle$ are the row and column average degrees, and N_r and N_c are the respective number of rows and columns nodes. The quantity is the sum of the term in equation (8) and a correction term taking into account the probability that links can appear in a row (column) also by sampling over the columns (rows). Specifically, the degree of a given row node is obtained by sampling over this axis (something which is correctly described by Equation (8)), or by sampling over columns: this information is embodied in the term within the square brackets proportional to k_j/N_r. In this latter case, one has to taken into account the probability that a links already exists because of the first sampling process, and this justifies the term proportional to $\langle k^r \rangle$. Of course, the final expression is symmetrical for row-column permutation.
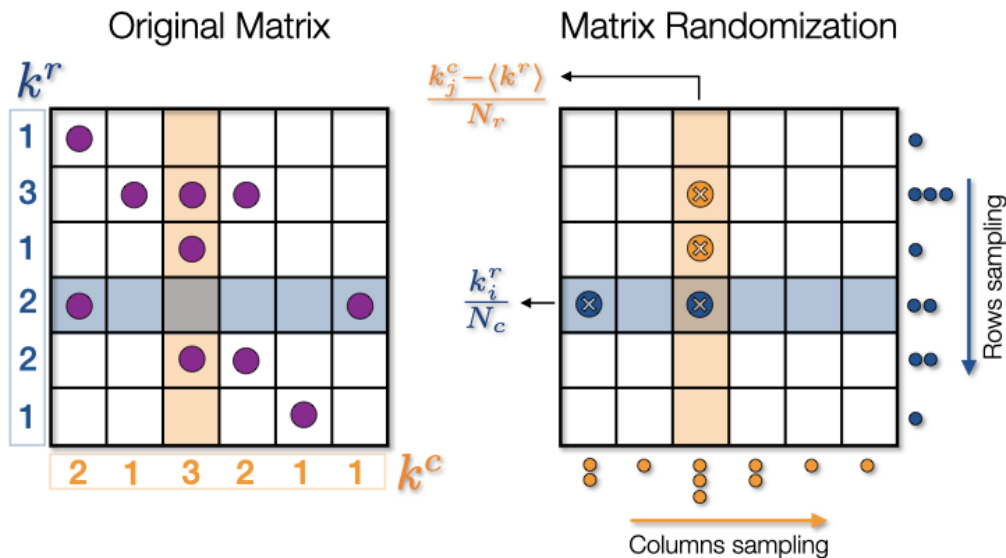

FIGURE 3 Corrected probabilistic model illustration. In the left we have the original matrix, while on the right the result of the null model randomization is shown. In this framework, the presence of a link is ruled by Equation (10), which extends Bascompte's model by taking into account the fact that row and column sampling processes are not independent. Particularly, the figure shows a situation where the row sampling introduces a link (in the fourth row) which affects the third column degree. Accordingly, the probability to have links in a given column is corrected by resting a quantity embodying the probability that a fraction of these links already exists because of the row sampling. In a first approximation, such an additional term may be reduced to the average row degree.

3 | RESULTS

In the following, we present the results concerning the test of the five null model classes against the three architectural patterns of interest. For a systematic exploration, Sections 3.1 and 3.2 exploit the synthetic generative model introduced in Sec. 2.2. With it, the construction of nested matrices is performed by keeping fixed the number of blocks ($B = 1$) and varying values of ξ and p, as in the top row of Fig. 2 . For the set of purely modular networks, we fixed $\xi = 1.5$ and $p = 1$ and smoothly modified B and μ to cover a wide range of modularity values, *e.g.* second row of Fig. 2 . Finally, the generation of in-block nested networks was carried out setting $B = 4$, $\xi = 1.5$ and varying levels of p and μ, *e.g.* third row of Fig. 2 . At the end of the process, we generated around 9000 networks for each one of the three patterns of interest, *i.e.*, $\sim 2.5 \times 10^4$ in total. For each generated network, we

created 100 random surrogates with each null model. The number of rows and columns were varied within the range $[30, 60]$. Finally, in Section 3.3, we perform the statistical significance test on a large collection of empirical networks.

3.1 | Resemblance of original *vs.* randomized networks

To understand the implications of each null model approach, we analyze in detail how a synthetic network resembles those of the corresponding null model ensembles. To do so, we first look into the node degrees of a particular matrix, and then perform a comparison over the space of parameters introduced in Sec. 2.2 by recalling the Jaccard distance.

Figure 4 analyzes in detail the properties of the ensemble obtained with the three PP models for a given network with planted nested structure (large matrix on the left). Examples for other types of matrices are available in Fig. S1 of the supplementary material. The top row represents the resulting matrices of interaction probability as a heatmap. Even visually, it is clear how permissive –or strict– each model is with respect to the original network. More in detail, the bottom row presents a scatter plot of the degree sequence of the original matrix against those of the resulting null matrices sets (in average). The error bars correspond to one standard deviation above and below the obtained average degrees. Herein, we observe that the probabilistic model appears to be too permissive, failing to preserve –even on average– the nodes' original degree. The max-entropy scheme almost exactly preserves the average degree sequences, hence, it might solve the proneness to type I errors that are present in the probabilistic model. The corrected probabilistic model presents a good agreement, with slight deviations for extreme (smallest and largest) degrees. Noteworthy, EE and FF models are not included in this figure –their scatter plots would render a trivial picture.
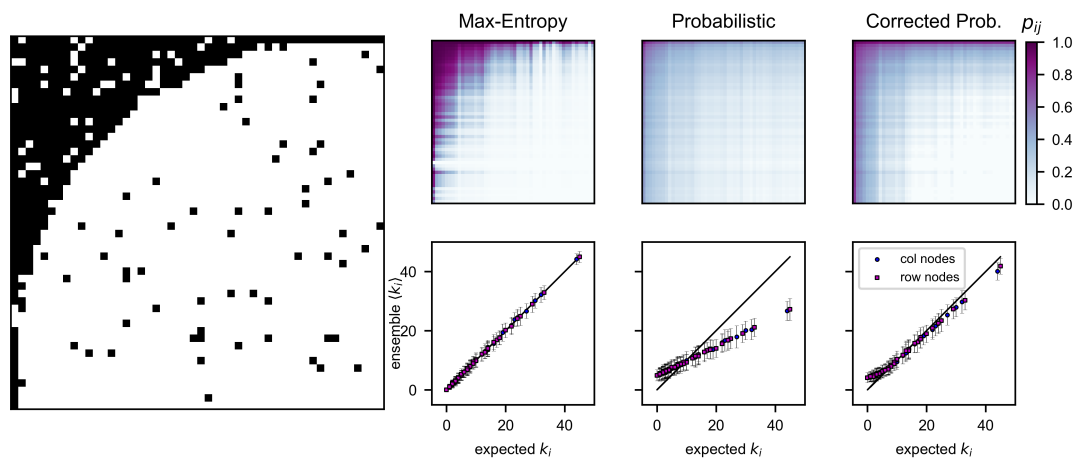

FIGURE 4 Comparison between a synthetic network (left) with planted nested structure, and their corresponding matrices of interaction probabilities p_{ij} (top right row) produced by the three PP models: maximum-entropy, Probabilistic and Corrected Probabilistic, respectively, and the scatter plots showing the fit between the obtained average degree sequences with each model and the real degrees of the matrix (bottom right row). Error bars in the scatter plots indicate one standard deviation above and below the average. The same comparisons between the three models for matrices with planted modular and in-block nested structures are available in Fig. S1 of the supplementary material.

For a wider look at the ensemble of null matrices, we now consider the parameter space of the network benchmark. Concretely, we explore to what extent the generated null matrices resemble the original synthetic ones for the whole range of parameters of the graph model. To quantify such resemblance, we employ the Jaccard distance J_d, which accounts for the level of dissimilarity between two binary arrays –the original and its random counterparts, in this case. It should be noted that, in the case of binary arrays, $J_d = 0$ indicates that the matrices are identical.

Top row in Fig. 5 shows the average Jaccard distance $\langle J_d \rangle$ over the nested $\xi - p$ parameter space for each null model scheme. Similar results for the modular and in-block nested matrices are shown in Fig. S2 of the supplementary material. As expected, the results for the EE and the probabilistic models show high values of the $\langle J_d \rangle$ over the entire $\xi - p$ parameter space, confirming that the randomized matrices generated by these models are substantially different from the original one. Turning to the FF and Max-entropy models, we observe a wide region of the $\xi - p$ parameter space in which $\langle J_d \rangle$ drops to intermediate and low

values –specially for networks with small noise p, regardless of the slimness of the nested structure ξ. These results for FF and Max-entropy confirm that, at least for nested networks, the resulting null matrices are very similar to the original one, since the models only allow for a very little deviation on the degree sequences –none in the FF model. Hence, these models might easily fall into type II errors. The corrected probabilistic model, instead, provides a good trade-off, offering a better fit of the degree sequences, while generating null matrices less similar to the original one, as evinced by the overall intermediate values of the $\langle J_d \rangle$ over the nested $\xi - p$ parameter space.
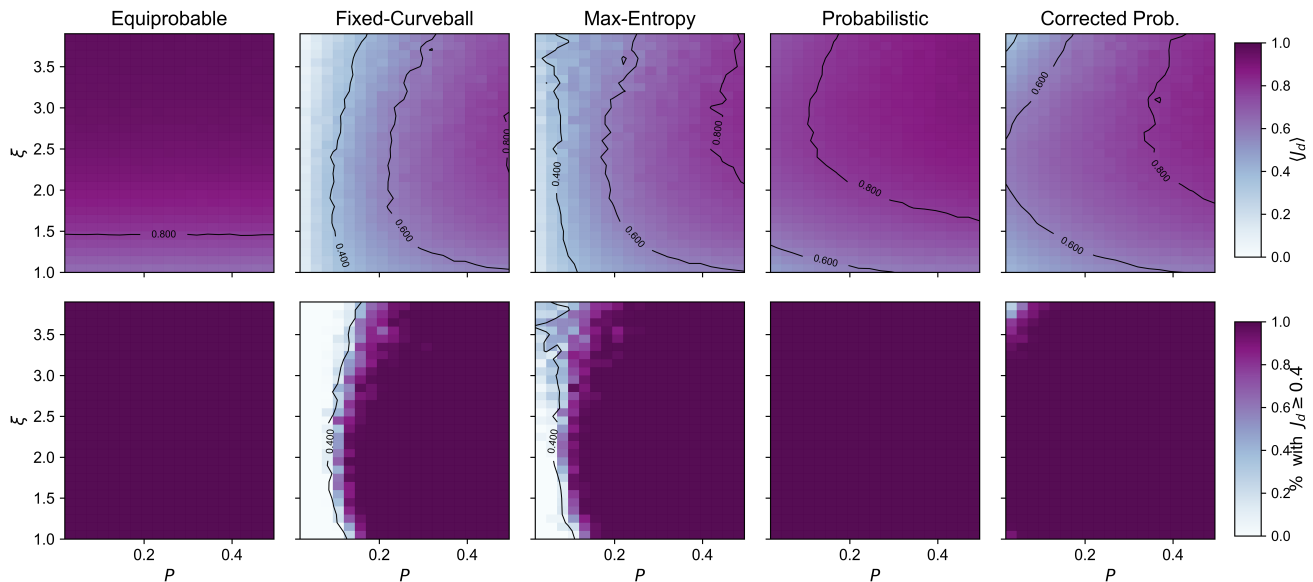

FIGURE 5 Synthetic nested Networks: $2-$dimensional plot in the $\xi - p$ parameter space showing: the average Jaccard distance $\langle J_d \rangle$ with of the null model ensembles with respect to the original nested synthetic matrix (top panel), and the % of matrices in the null model ensembles for which the Jaccard distance with respect to the original synthetic matrix was above a threshold of $J_d \geq 0.4$ (bottom panel), for the five models studied in this work.

From another perspective, the bottom panel of Fig. 5 shows the percentage of matrices in the null model ensembles for which $J_d \geq 0.4$ in the nested $\xi - p$ parameter space. Clearly, there are relevant regions for which the FF and the Max-entropy models cannot generate true variants of the original matrix, even with such moderate threshold. Of course, when imposing a more rigid threshold this region becomes wider. In the other extreme, the EE and Probabilistic schemes prove to have ample room to generate largely different networks, no matter the parameter values under study. Somewhere in between, the Corrected Probabilistic model is flexible enough to produce different networks, except in some extreme cases (very sparse networks with low p). These insights are all in good agreement with a recent work developed by Strona *et al.*[74], in which they proved that slightly relaxing the restrictiveness in the preservation of the rows and column totals –quasi-FF matrices– can reduce the tendency to underestimate significant nestedness.

3.2 | Statistical significance of structural patterns

We now analyze how consistent the five null models described above are in assessing the statistical significance of nested, modular and in-block nested patterns, for each set of the generated synthetic matrices (192 empirical ecological networks are analyzed in the next section). In each case, we assessed the significance of the three patterns of interest. The results of the analysis are summarized in Table 1 , which accounts the % of networks with $z \geq 2$, *i.e.*, an associated p-value around 0.025. For nested networks, we also explore the behavior of the null models under other nestedness metrics: $NODF$[54] and the spectral radius $\tilde{\rho}$[51,11], normalized with respect to the square root of the number of links in the matrices, $\sqrt{L}$, as in[32].

For networks with planted nested structure ($B = 1$, $p \in [0, 0.5]$), the EE and Probabilistic models overestimate systematically the statistical significance for nestedness –no matter the metric. In the other extreme, both the FF and Max-entropy restrictive

Planted structure		EE	FF	PP		
		Equiprobable	Fixed Curveball	Max-entropy	Probabilistic	Prob. Corrected
Nested	$\mathcal{N}$	100	1.28	5.9	99.88	41.33
	$NODF$	100	0.06	4.73	99.93	46.21
	$\tilde{\rho}$	100	28.13	3.48	99.02	58.39
	Q	6.24	24.2	46.37	11.28	19.32
	$\mathcal{I}$	87.53	4.04	7.33	85.84	44.65
Modular	$\mathcal{N}$	0.76	1.63	0.	0.	0.
	Q	72.96	72.34	85.99	70.74	81.12
	$\mathcal{I}$	50.93	52.94	40.21	44.57	38.63
In-block nested	$\mathcal{N}$	74.81	32.76	0	28.28	0.05
	Q	56.82	87.64	97.1	84.71	92.86
	$\mathcal{I}$	55.89	69.75	67.99	69.05	63.22

TABLE 1 Percentage % of networks with $z \geq 2$ for three sets of synthetic networks with nested, modular and in-block nested structure. The over and underestimation values mentioned in the text are marked in red and blue, respectively.

schemes deliver an underestimation of the amount of truly nested networks. Thus, confirming their proneness to type I and II errors, respectively. Only the corrected probabilistic model renders a balanced assessment, consistent with the parameters used in the generation of this set of nested synthetic matrices. The assessment of the significance for in-block nestedness is generally consistent with the one observed for nestedness, as $\mathcal{I}$ reduces to $\mathcal{I}$ when $B = 1$ (recall Eqs. 2 and 4).

More intriguingly, the FF and Max-entropy models identify a large fraction of nested-by-construction networks as being significantly modular (24.2% and 46.37%, respectively). To some extent, this could be explained by the densest nested matrices in the collection[38], *i.e.*, low ξ, but by no means the collection should be regarded, overall, to be more modular than nested.

Turning to modular-by-construction networks ($B > 1$, $p = 1$, $\mu \in [0, 0.5]$), we observe that –unlike the previous case– all models render intuitively correct results: these networks are unanimously assessed as non-nested, mostly assessed as modular (except for those with high μ), and around half of the collection is assessed as being in-block nested, which makes sense considering configurations with low p (which are modular *and* in-block nested).

Finally, networks with planted in-block nestedness ($B = 4$, $p \in [0, 0.5]$, $\mu \in [0, 0.5]$) again arise inconsistencies among the randomization procedures. Some of them (EE, FF and Probabilistic) incorrectly assess a large amount of networks as being nested –this is particularly marked in the case of EE.

Considering the results in Table 1 , it is clear that nested structures are the most confusing given the outcomes of the different randomization schemes. For this reason, we explore these inconsistencies in further detail in Figure 6 regarding only the nested-by-construction synthetic subset. The figure shows the z-score values for nestedness $z_{\mathcal{N}}$ (top) and modularity z_Q (bottom) over the $\xi - p$ parameter space of the generated nested matrices, considering the three PP models. Thumbnails of the resulting generated matrices over the parameter space are shown on the left part of the figure as a visual aid.

The results of Fig. 6 for the probabilistic model confirm the tendency for overestimating the statistical significance of nested patterns, *i.e.*, type I error. The model yields values of $z_{\mathcal{N}} \geq 2$ over the whole parameter space, including the regions where the planted nested structure becomes hardly identifiable (high p). The Max-entropy model, conversely, is not able to assess the significance of nestedness, even in regions where the networks have a clearly defined nested pattern, except for those networks with high ξ. The corrected probabilistic model is capable of assessing the significance of the nested pattern over a wide region of the parameter space, for intermediate to low values of noise p –as expected–, and different shape values ξ. Note, however, that the higher the value of ξ is, less noise p is necessary for the model to reach $z_{\mathcal{N}} \leq 2$. This is not surprising, since the ξ parameter affects the network connectance, and sparser networks will require a lower level of noise to distort the planted nested structure. For the sake of completeness, Fig. S3 in the supplementary material shows the results for $NODF$ and $\tilde{\rho}$, showing equivalent results, in spite of the different nature of the two metrics. The assessment of significance for modularity z_Q with the probabilistic and corrected probabilistic models shows small regions of $z_Q \geq 2$ that corresponds to high connectance networks (low ξ). In these cases, one can assume that networks are dense enough, such that the optimization algorithm employed for modularity is able to find a partition with $B > 1$, resulting in values of Q larger than expected[38]. The Max-entropy model, however, assesses
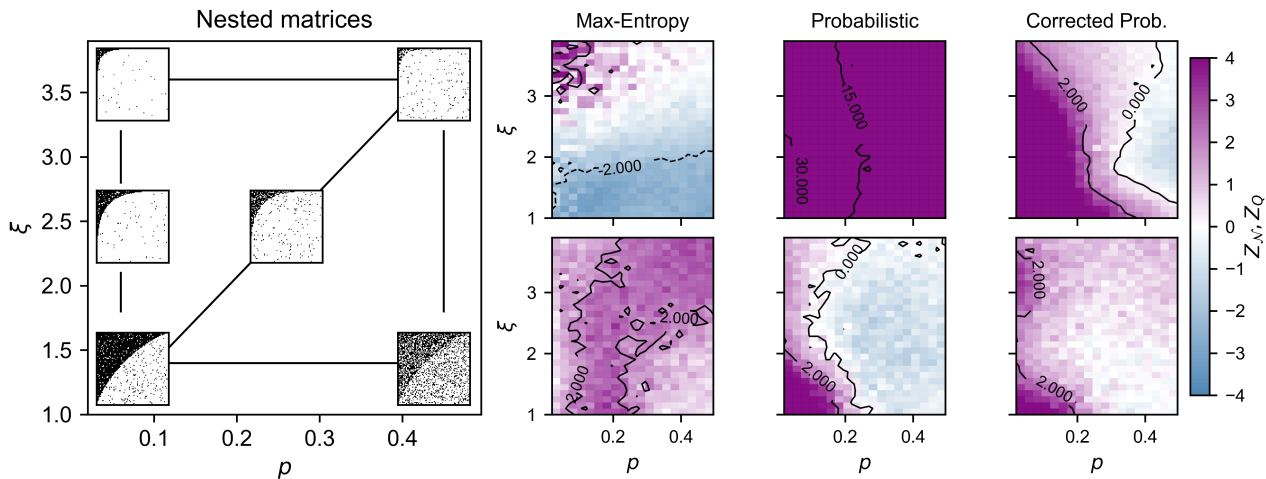
FIGURE 6 Synthetic Nested Networks: Heatmaps in the $\xi - p$ parameter plane showing: $z_{\mathcal{N}}$ (top) and z_Q (bottom) scores values under the Max-entropy, probabilistic, and corrected probabilistic models for the set of ~9000 synthetic bipartite networks with planted nested structure. Examples of the resulting generated matrices over the parameter space are shown on the left part of the figure.

significance for modular patterns over a wide region of the $\xi - p$ parameter space –around 46% of the networks in Table 1 –, strengthening our assumptions regarding the presence of biases towards an overestimation of statistical significance of modular patterns.
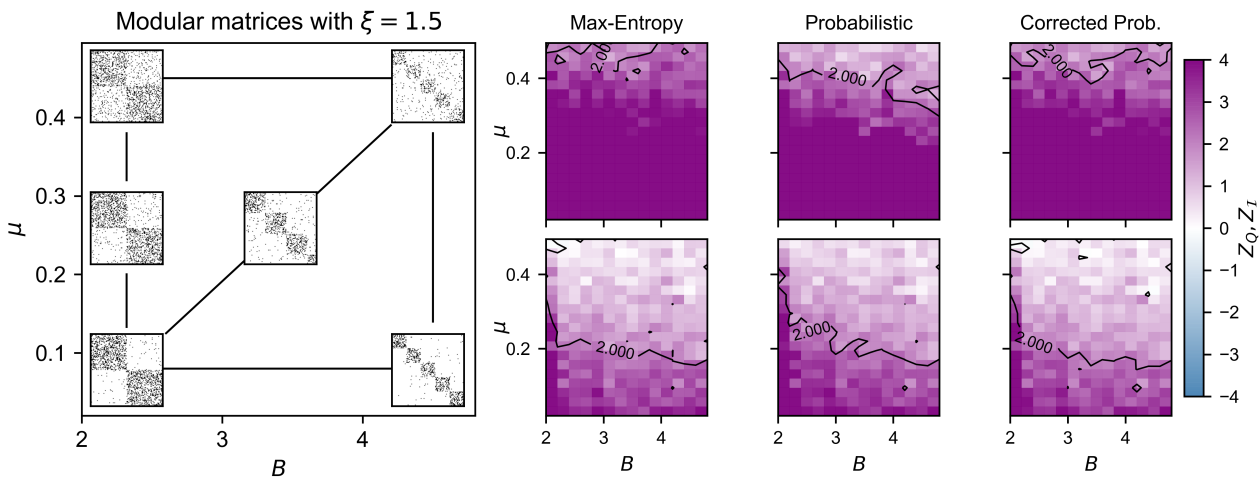

FIGURE 7 Synthetic Modular Networks: Heatmaps in the $\mu - B$ parameter plane showing: z_Q (top) and z_I (bottom) scores values under the Max-entropy, probabilistic, and corrected probabilistic models for the set of ~9000 synthetic bipartite networks with modular structure. Examples of the resulting generated matrices over the parameter space are shown on the left part of the figure.

Moving on, we know discuss the results for the set of synthetic matrices with modular structure. From Table 1 , we can observe that the assessment of significance for nestedness across this set of matrices yields coherent results for the five models. We obtained values of $z_{\mathcal{N}} \leq 2$ consistently throughout the whole set. Such result is not surprising, since by construction, all the networks present a clearly modular planted structure, and as mentioned above, we have evidence that these two patterns may not be structurally compatible[65]. Results for the assessment of modularity and in-block nestedness from Table 2 are complemented with Figure 7 . The figure shows the z-score values for modularity z_Q (top) and in-block nestedness z_I (bottom) over the $\mu - B$

parameter space of the generated matrices, with the three PP models. Similarly to Fig. 6 , examples of the resulting matrices over the parameter space are shown on the left part of the figure. Again, Table 1 reports an unanimous assessment of modularity across the five models. Top panel in Fig.7 for the PP models confirms these results: the models are able to assess the significance of modularity for a wide range of the parameter space. As expected, the assessment is sensitive to level inter-block noise μ. Although, the Max-entropy model is still able to assess the significance for modularity in regions with high μ values, where the planted modular structure is hardly identifiable anymore. Results for the assessment of in-block nestedness also show consistent results across the five models (see Table 1). For around half of the networks in this set, we obtain a significant level of in-block nestedness. Again, the critical ingredient here is the presence of inter-block noise μ. The result is in agreement with Palazzi *et al.* [65], in which it is proved that both Q and $\mathcal{I}$ can coexist, as there are no mutually imposed constraints between them.
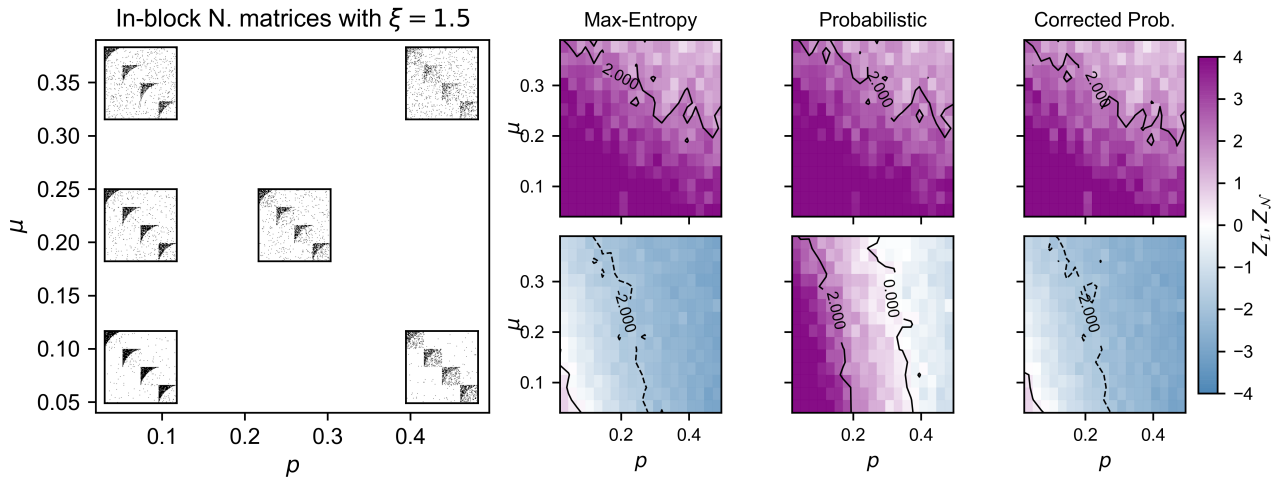

FIGURE 8 Synthetic in-block nested Networks: Heatmaps in the $\mu - p$ parameter plane showing: $z_{\mathcal{I}}$ (top) and $z_{\mathcal{N}}$ (bottom) scores values under the Max-entropy, probabilistic, and corrected probabilistic models for the set of $\sim$9000 synthetic bipartite networks with in-block nested structure. Examples of the resulting generated matrices over the parameter space are shown on the left part of the figure.

Finally, we move on to examine the results for the last set of synthetic networks, with planted in-block nested structure. Figure 8 complements the results for this set of networks in Table 1 , by showing the z-score values for in-block nestedness $z_{\mathcal{I}}$ (top) and nestedness $z_{\mathcal{N}}$ (bottom). As in the previous cases, the left part of the figure shows examples of the generated matrices with the PP models over the $\mu - p$ parameter space. The three models illustrated here show similar (and expected) results for $z_{\mathcal{I}}$: in-block nestedness becomes less significant when the structure is distorted within (higher p) or among (higher μ) blocks. Such coherence contrasts with the assessment of significance for $\mathcal{N}$, which deserve to be examined in more detail. It is known that global nestedness decreases with the number of blocks in a network, regardless of their internal nested structure [65]. Consequentially, the Max-entropy and Corrected probabilistic models correctly present non-significant $z_{\mathcal{N}} < 2$ throughout the parameter plane, and even for low values of noise (either μ or p). Instead, the probabilistic scheme identifies as nested those matrices for which $p < 0.2$, regardless of μ, a clear overestimation of nested patterns which occurs, even at higher proportions, in the EE and FF models (see Table 1).

3.3 | Empirical Networks

Lastly, we also analyze how consistent the five null models are in assessing the statistical significance of nested, modular, and in-block nested patterns over a set of 192 empirical networks. All networks are treated as binary, and that only networks between 20 and 300 nodes are kept (see Section S2 and Figure S4 in the supplementary material for details). For the sake of consistency and completeness, we explore the behavior of the null models under the three nestedness metrics: $\mathcal{N}$, NODF, and $\tilde{\rho}$. We report the percentage of networks with $z \geq 2$ (p-value ~ 0.025) for each one of our five structural metrics in Table 2 .

Assessment of nestedness in the empirical data for the EE and Probabilistic models is consistent across the three different metrics, yielding z-score values $z \geq 2$ for a large proportion of the analyzed networks. The assessment of significance for in-block nestedness renders equivalent results, showing significant $z_{\mathcal{I}\geq}$ for a high percentage of the networks, although in a lower proportion with respect to nestedness. In contrast, only a small fraction of the networks is significantly modular ($z_Q \geq 2$). Results for the strictest schemes –FF and Max-entropy models–, show, in the case of nestedness, a good agreement with the results in the synthetic matrices (and the literature [72]), yielding non-significant z-scores for a large majority of the networks. Note that, similar to the synthetic case, the normalized spectral radius $\tilde{\rho}$, with the FF model is able to assess nestedness for a bigger proportion of the networks. Once again, the corrected probabilistic model seems to provide a more balanced assessment of nested patterns over the different metrics. The small variations in the preservation of the node's degrees in the corrected probabilistic model allow nestedness to reemerge as significant in a small proportion of the networks.

| | | EE | FF | PP | | |
		Equiprobable	Fixed Curveball	Max-entropy	Probabilistic	Corrected prob.
Empirical	$\mathcal{N}$	80.73	2.08	0.52	67.19	13.02
	$NODF$	83.33	1.04	10.94	69.27	29.69
	$\tilde{\rho}$	87.5	22.9	2.08	62.5	20.31
	Q	10.42	14.06	66.67	15.62	36.98
	$\mathcal{I}$	52.6	20.31	32.29	46.35	34.38

TABLE 2 Percentage of networks with score $z \geq 2$ for a collection of 192 empirical ecological networks.

Moving on, results for modularity in the FF and Max-entropy model render inconsistent results, suggesting once again the presence of biases in the Max-entropy model with regard to the assessment of modular patterns. Finally, we observe that a relevant fraction of the networks exhibit significant in-block nestedness structure even with the most restrictive algorithms, in close agreement with recent studies suggesting the presence of in-block nestedness as a frequent pattern in ecological communities [42,40,39].

4 | DISCUSSION

The overwhelming evidence that real ecological networks are highly structured, rather than randomly assembled, has sparked much research to explain the emergence and dynamical effects of those interaction patterns. In parallel, efforts have been devoted at determining the statistical significance of such patterns. As a result, the use and design of specialized null models to assess the statistical significance of selected architectures has become a widespread approach within the ecological community. However, to this day, there is still no agreement on which is the most suitable procedure to perform this kind of analysis. The discussion regarding the selection on the level of constraints to build the null matrices, the construction of synthetic benchmarks, or the appropriate null model-metric combination remains an open issue among scholars. Furthermore, there is already abundant evidence suggesting that networks often exhibit hybrid structures and/or at different scales of organization, yet the concurrent assessment of the statistical significance of such complex patterns has not been extensively studied.

In this work, we have first addressed the need to build synthetic networks with total control over their dominant structure. Our generative benchmark allows for the construction of networks with clear defined architectural patterns, transitioning from structured (nested, modular, in-block nested) to random networks in a smooth way. Only then we can generate a rich sample to systematically study of the performance of five popular null models.

With over 25000 networks of varying size, connectance and internal structure, we have observed the existence of ambiguities and biases in the concurrent assessment of network patterns, either over- or underestimating their significance. In particular, our use of an extensive ensemble of networks uncovers the weaknesses of both too restrictive (FF, Max-entropy) and too permissive (EE, Probabilistic) randomizing procedures. At the face of such results, we propose a variation of the Probabilistic model, that mitigates some of the limitations exhibited by the current models –in close agreement with recent research, regarding the convenience of quasi-FF approaches [74]. Last but not least, our results on empirical networks show an overall good agreement

with the results obtained on the synthetic matrices. Indeed, we observe again the presence of ambiguities in the simultaneous assessment of the patterns across models, with the corrected probabilistic model standing out as a reliable alternative to assess the significance of highly structured matrices.

We are convinced that our results can orient the future debate regarding the use of null models in bipartite networks within and beyond the ecological community. Providing a compact framework and off-the-shelf implementations (synthetic benchmark, structural descriptors, optimization algorithms, and null models), our aim is double: on the theoretical side, to continue the research on the underlying assumptions and implicit consequences of the different null models. For example, to discern the apparent bias of the Max-entropy model towards underestimating the presence of nested configurations, while overestimating the presence of modular ones. On the practical side, to encourage researchers to shift the focus towards the simultaneous assessment of the statistical significance of multiple structural patterns at multiple scales –especially given the increasing evidence that empirical systems may display structure at several scales of organization, and frequently in the form of compound structures.

CONFLICT OF INTERESTS

The authors declare that there are no conflicts of interest regarding the publication of this paper.

CODE AVAILABILITY

The corresponding software code to (1) generate synthetic networks with nested, modular and in-block nested structures; (2) the implementation of the metrics and optimization algorithms aimed to characterize nestedness, modularity, and in-block nestedness; and (3) the Python implementation of the five null models employed in this paper are available at the Github repository https://github.com/COSIN3-UOC.

ACKNOWLEDGEMENTS

References

1. Bascompte J, Jordano P, Melián CJ, Olesen JM. The nested assembly of plant–animal mutualistic networks. *Proceedings of the National Academy of Sciences* 2003; 100(16): 9383–9387.

2. Olesen JM, Bascompte J, Dupont YL, Jordano P. The modularity of pollination networks. *Proceedings of the National Academy of Sciences* 2007; 104(50): 19891–19896.

3. Bastolla U, Fortuna MA, Pascual-García A, Ferrera A, Luque B, Bascompte J. The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature* 2009; 458(7241): 1018–1020.

4. Suweis S, Simini F, Banavar JR, Maritan A. Emergence of structural and dynamical properties of ecological mutualistic networks. *Nature* 2013; 500(7463): 449.

5. Allesina S, Grilli J, Barabás G, Tang S, Aljadeff J, Maritan A. Predicting the stability of large structured food webs. *Nature Communications* 2015; 6(1): 7842.

6. Grilli J, Rogers T, Allesina S. Modularity and stability in ecological communities. *Nature Communications* 2016; 7(1): 1–10.

7. Okuyama T, Holland JN. Network structural properties mediate the stability of mutualistic communities. *Ecology Letters* 2008; 11(3): 208-216. doi: 10.1111/j.1461-0248.2007.01137.x

8. Thébault E, Fontaine C. Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science* 2010; 329(5993): 853–856.

9. Allesina S, Tang S. Stability criteria for complex ecosystems. *Nature* 2012; 483(7388): 205.

10. Stouffer DB, Bascompte J. Compartmentalization increases food-web persistence. *Proceedings of the National Academy of Sciences* 2011; 108(9): 3648–3652. doi: 10.1073/pnas.1014353108

11. Staniczenko PP, Kopp JC, Allesina S. The ghost of nestedness in ecological networks. *Nature Communications* 2013; 4: 1391.

12. Rohr RP, Saavedra S, Bascompte J. On the structural stability of mutualistic systems. *Science* 2014; 345(6195): 1253497.

13. Saavedra S, Rohr R, Olesen J, Bascompte J. Nested species interactions promote feasibility over stability during the assembly of a pollinator community. *Ecology and Evolution* 2015; 6(4): 1007.

14. Pascual-García A, Bastolla U. Mutualism supports biodiversity when the direct competition is weak. *Nature Communications* 2017; 8(1): 1–13.

15. Mariani MS, Ren ZM, Bascompte J, Tessone CJ. Nestedness in complex networks: Observation, emergence, and implications. *Physics Reports* 2019; 813: 1-90. Nestedness in complex networks: Observation, emergence, and implicationsdoi: https://doi.org/10.1016/j.physrep.2019.04.001

16. Ulrich W, Gotelli NJ. Null model analysis of species nestedness patterns. *Ecology* 2007; 88(7): 1824–1831.

17. Ulrich W, Almeida-Neto M, Gotelli NJ. A consumer's guide to nestedness analysis. *Oikos* 2009; 118(1): 3–17.

18. Gotelli NJ, Graves GR. *Null models in ecology*. Smithsonian Institution Press. . 1996.

19. Gotelli NJ, Ulrich W. Statistical challenges in null model analysis. *Oikos* 2012; 121(2): 171–180.

20. Beckett SJ, Boulton CA, Williams HT. FALCON: a software package for analysis of nestedness in bipartite networks. *F1000Research* 2014; 3.

21. Strona G, Fattorini S. On the methods to assess significance in nestedness analyses. *Theory in Biosciences* 2014; 133(3-4): 179–186.

22. Gotelli NJ, Entsminger GL. Swap and fill algorithms in null model analysis: rethinking the knight's tour. *Oecologia* 2001; 129(2): 281–291.

23. Sale PF. Overlap in resource use, and interspecific competition. *Oecologia* 1974; 17(3): 245–256.

24. Connor EF, Simberloff D. The assembly of species communities: chance or competition?. *Ecology* 1979; 60(6): 1132–1140.

25. Miklós I, Podani J. Randomization of presence–absence matrices: comments and new algorithms. *Ecology* 2004; 85(1): 86–92.

26. Strona G, Nappo D, Boccacci F, Fattorini S, San-Miguel-Ayanz J. A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals. *Nature Communications* 2014; 5(1): 1–9.

27. Miller JW, Harrison MT. Exact sampling and counting for fixed-margin matrices. *The Annals of Statistics* 2013; 41(3): 1569–1592.

28. Rechner S, Strowick L, Müller-Hannemann M. Uniform sampling of bipartite graphs with degrees in prescribed intervals. *Journal of Complex Networks* 2018; 6(6): 833–858.

29. Artzy-Randrup Y, Stone L. Generating uniformly distributed random networks. *Physical Review E* 2005; 72(5): 056708.

30. Gotelli NJ. Null model analysis of species co-occurrence patterns. *Ecology* 2000; 81(9): 2606–2621.

31. Dormann CF, Fründ J, Blüthgen N, Gruber B. Indices, graphs and null models: analyzing bipartite ecological networks. *The Open Ecology Journal* 2009; 2(1).

32. Bruno M, Saracco F, Garlaschelli D, Tessone CJ, Caldarelli G. The ambiguity of nestedness under soft and hard constraints. *Scientific Reports* 2020; 10(1): 1–13.

33. Beckett S, Williams H. Synthetic matrix ensemble for nestedness analysis. 2015

34. Fortuna MA, Stouffer DB, Olesen JM, et al. Nestedness versus modularity in ecological networks: two sides of the same coin?. *Journal of Animal Ecology* 2010; 79(4): 811–817.

35. Thébault E. Identifying compartments in presence–absence matrices and bipartite networks: insights into modularity measures. *Journal of Biogeography* 2013; 40(4): 759–768.

36. Lewinsohn TM, Inácio Prado P, Jordano P, Bascompte J, Olesen JM. Structure in plant–animal interaction assemblages. *Oikos* 2006; 113(1): 174–184.

37. Flores CO, Meyer JR, Valverde S, Farr L, Weitz JS. Statistical structure of host–phage interactions. *Proceedings of the National Academy of Sciences* 2011; 108(28): E288–E297.

38. Solé-Ribalta A, Tessone CJ, Mariani MS, Borge-Holthoefer J. Revealing In-Block Nestedness: detection and benchmarking. *Physical Review E* 2018; 96(6): 062302.

39. Lampo A, Palazzi MJ, Borge-Holthoefer J, Solé-Ribalta A. Hybrid structural arrangements mediate stability and feasibility in mutualistic networks. *https://arxiv.org/pdf/2105.09191.pdf* 2021.

40. Cai W, Snyder J, Hastings A, D'Souza RM. Mutualistic networks emerging from adaptive niche-based interactions. *Nature Communications* 2020; 11(1). doi: https://doi.org/10.1038/s41467-020-19154-5

41. Palazzi MJ, Solé-Ribalta A, Calleja-Solanas V, et al. An ecological approach to structural flexibility in online communication systems. *Nature Communications* 2021; 12(1): 1–11.

42. Mello MA, Felix GM, Pinheiro RB, et al. Insights into the assembly rules of a continent-wide multilayer network. *Nature Ecology & Evolution* 2019: 1–8.

43. Hultén E. *Outline of the history of arctic and boreal biota during the Quaternary period.* Bokforlags Aktiebolaget Thule . 1937.

44. Darlington PJ. Zoogeography. *The Geographical Distributions of Animals* 1957.

45. Patterson BD, Atmar W. Nested subsets and the structure of insular mammalian faunas and archipelagos. *Biological Journal of the Linnean Society* 1986; 28(1-2): 65–82.

46. Soramäki K, Bech ML, Arnold J, Glass RJ, Beyeler WE. The topology of interbank payment flows. *Physica A: Statistical Mechanics and its Applications* 2007; 379(1): 317-333. doi: https://doi.org/10.1016/j.physa.2006.11.093

47. König MD, Tessone CJ, Zenou Y. Nestedness in networks: A theoretical model and some applications. *Theoretical Economics* 2014; 9(3): 695-752. doi: https://doi.org/10.3982/TE1348

48. Saracco F, Di Clemente R, Gabrielli A, Squartini T. Detecting early signs of the 2007–2008 crisis in the world trade. *Scientific Reports* 2016; 6(1): 30286. doi: https://doi.org/10.1038/srep30286

49. Borge-Holthoefer J, Baños RA, Gracia-Lázaro C, Moreno Y. Emergence of consensus as a modular-to-nested transition in communication dynamics. *Scientific Reports* 2017; 7: 41673.

50. Payrató-Borràs C, Hernández L, Moreno Y. Measuring nestedness: A comparative study of the performance of different metrics. *Ecology and Evolution* 2020; 10(21): 11906–11921.

51. Bell FK, Cvetkovic D, Rowlinson P, Simic SK. Graphs for which the least eigenvalue is minimal, II. *Linear Algebra and its Applications* 2008; 429(8): 2168-2179. doi: https://doi.org/10.1016/j.laa.2008.06.018

52. Bhattacharya A, Friedland S, Peled UN. On the first eigenvalue of bipartite graphs. *arXiv preprint arXiv:0809.1615* 2008.

53. Atmar W, Patterson BD. The measure of order and disorder in the distribution of species in fragmented habitat. *Oecologia* 1993; 96(3): 373–382.

54. Almeida-Neto M, Guimaraes P, Guimarães PR, Loyola RD, Ulrich W. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos* 2008; 117(8): 1227–1239.

55. Galeano J, Pastor JM, Iriondo JM. Weighted-interaction nestedness estimator (WINE): a new estimator to calculate over frequency matrices. *Environmental Modelling & Software* 2009; 24(11): 1342–1346.

56. Almeida-Neto M, Ulrich W. A straightforward computational approach for measuring nestedness using quantitative matrices. *Environmental Modelling & Software* 2011; 26(2): 173–178.

57. Podani J, Ricotta C, Schmera D. A general framework for analyzing beta diversity, nestedness and related community-level phenomena based on abundance data. *Ecological Complexity* 2013; 15: 52–61.

58. Zachary WW. An information flow model for conflict and fission in small groups. *Journal of Anthropological Research* 1977; 33(4): 452–473.

59. Guimerà R, Amaral LAN. Functional cartography of complex metabolic networks. *Nature* 2005; 433: 895-900.

60. Adamic LA, Glance N. The political blogosphere and the 2004 US election: divided they blog. In: ACM. ; 2005: 36–43.

61. Fortunato S. Community detection in graphs. *Physics Reports* 2010; 486(3-5): 75–174.

62. Borge-Holthoefer J, Arenas A. Semantic Networks: Structure and Dynamics. *Entropy* 2010; 12(5): 1264–1302. doi: 10.3390/e12051264

63. Fortunato S, Hric D. Community detection in networks: A user guide. *Physics Reports* 2016; 659: 1–44. doi: 10.1016/j.physrep.2016.09.002

64. Newman ME, Girvan M. Finding and evaluating community structure in networks. *Physical Review E* 2004; 69(2): 026113.

65. Palazzi M, Borge-Holthoefer J, Tessone C, Solé-Ribalta A. Macro-and mesoscale pattern interdependencies in complex networks. *Journal of the Royal Society Interface* 2019; 16(159): 20190553.

66. Palazzi MJ, Cabot J, Izquierdo JLC, Solé-Ribalta A, Borge-Holthoefer J. Online division of labour: emergent structures in Open Source Software. *Scientific Reports* 2019; 9(1): 1–11.

67. Flores CO, Valverde S, Weitz JS. Multi-scale structure and geographic drivers of cross-infection within marine bacteria and phages. *The ISME Journal* 2013; 7(3): 520–532.

68. Kemp JE, Evans DM, Augustyn WJ, Ellis AG. Invariant antagonistic network structure despite high spatial and temporal turnover of interactions. *Ecography* 2017; 40(11): 1315–1324.

69. Guardiola M, Stefanescu C, Rodà F, Pino J. Do asynchronies in extinction debt affect the structure of trophic networks? A case study of antagonistic butterfly larvae–plant networks. *Oikos* 2018; 127(6): 803–813.

70. Web of Life: ecological networks database. http://www.web-of-life.es/; 2012.

71. Gilpin ME, Diamond JM. Factors contributing to non-randomness in species co-occurrences on islands. *Oecologia* 1982; 52(1): 75–84.

72. Payrató-Borràs C, Hernández L, Moreno Y. Breaking the spell of nestedness: The entropic origin of nestedness in mutualistic systems. *Physical Review X* 2019; 9(3): 031024.

73. Joppa LN, Montoya JM, Sanderson J, Pimm SL, others . On nestedness in ecological networks. *Evolutionary Ecology Research. 2010; 12: 35-46* 2010.

74. Strona G, Ulrich W, Gotelli NJ. Bi-dimensional null model analysis of presence-absence binary matrices. *Ecology* 2018; 99(1): 103–115.

75. Lehsten V, Harmand P. Null models for species co-occurrence patterns: assessing bias and minimum iteration number for the sequential swap. *Ecography* 2006; 29(5): 786–792.

76. Carstens CJ. Proof of uniform sampling of binary matrices with fixed row sums and column sums for the fast Curveball algorithm. *Physical Review E* 2015; 91(4): 042812.

77. Squartini T, Garlaschelli D. Analytical maximum-likelihood method to detect patterns in real networks. *New Journal of Physics* 2011; 13(8): 083001.

78. Saracco F, Di Clemente R, Gabrielli A, Squartini T. Randomizing bipartite networks: the case of the World Trade Web. *Scientific Reports* 2015; 5(1): 1–18.

How to cite this article: M. J. Palazzi, A. Lampo, A Solé-Ribalta, and J. Borge-Holthoefer (2021), Understanding and overcoming biases in the use of null models for bipartite networks, *Methods in Ecology and Evolution*, *2021;xx:x–x*.