Revealing In-Block Nestedness: detection and benchmarking

Albert Solé-Ribalta  
*Internet Interdisciplinary Institute (IN3), Universitat Oberta de Catalunya, Barcelona, Catalonia, Spain*

Claudio J. Tessone  
*URPP Social Networks, Universität Zürich, Switzerland*

Manuel S. Mariani  
*Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, PR China  
URPP Social Networks, Universität Zürich, Switzerland and  
Physics Department, Université de Fribourg, Switzerland*

Javier Borge-Holthoefer  
*Internet Interdisciplinary Institute (IN3), Universitat Oberta de Catalunya, Barcelona, Catalonia, Spain and  
Institute for Biocomputation and Physics of Complex Systems (BIFI), Universidad de Zaragoza, Zaragoza, Spain*  
(Dated: January 18, 2018)

As new instances of nested organization—beyond ecological networks—are discovered, scholars are debating around the co-existence of two apparently incompatible macroscale architectures: nestedness and modularity. The discussion is far from being solved, mainly for two reasons. First, nestedness and modularity appear to emerge from two contradictory dynamics, cooperation and competition. Second, existing methods to assess the presence of nestedness and modularity are flawed when it comes to the evaluation of concurrently nested and modular structures. In this work, we tackle the latter problem, presenting the concept of *in-block nestedness*, a structural property determining to what extent a network is composed of blocks whose internal connectivity exhibits nestedness. We then put forward a set of optimization methods that allow us to identify such organization successfully, both in synthetic and in a large number of real networks. These findings challenge our understanding of the topology of ecological and social systems, calling for new models to explain how such patterns emerge.

I. INTRODUCTION

The identification of macroscale connectivity patterns has been central to the development of network science. Beyond the inherent methodological challenges of this task, ascertaining them is of relevance to the specific disciplines and to the area as a whole, inasmuch they are the outcome of distinct microscopic mechanisms of network formation. It is in this context—i.e. understanding network architecture as an emergent feature—that nestedness and modularity arise as prominent macro-structural signatures to study.

The concept of nestedness was first coined in biology to characterize the spatial distribution of biotas in isolated, yet spatially-related landscapes [1], and later found to describe large families of inter-species cooperative relations [2]. In structural terms, a perfectly nested pattern is such that the set of connections of any given node is a subset of the relationships of larger degree ones [3]; see Fig. 1 (left). Nestedness has imposed itself as a landmark feature—nestedness and modularity arise as prominent macro-structural signatures to study.

Beyond natural systems, nestedness emerges as well in social, technical, and economic systems, e.g. industrial relationships [4,6,7], international trade [8], information ecosystems [9], anthropology [10] and knowledge production [11]. In socio-economic systems, epitome of this property in unipartite networks, emergence of nestedness is originated in agents attempting to maximize their own centrality [12,13].

On the other side, the identification of modular patterns in networks stands as one of the hallmarks in the area with prominent precedents in social network analysis [15]. Besides social systems, networks with significant community structure, see Fig. 1 (middle), appear in multiple contexts [16], like biology [17] or cognitive science [18]. It implies the existence of subgroups of nodes, strongly connected within but loosely connected to nodes outside. The identification and analysis of community structure constitutes itself a sub-area of network science. It poses challenges with respect to detection algorithms, empirical problems, applications and conclusions derived [19].

Nestedness and modularity have been often treated as incompatible architectures, since they are thought to emerge from conflicting (respectively, cooperative and competitive) dynamics [20]. Thus, most studies have focused exclusively on either of them. The existence of systems which combine both patterns has been largely overlooked, despite challenging indications in natural [21,24] and social ecosystems [9]. As of now, the proper identification of such compound structures lays beyond the capabilities of state-of-the-art techniques.

In the scarce existing literature we identify two different approaches. The first operates in parallel, measur-
The paper is organized as follows: to overcome the limitations of existing approaches, Section III introduces a compact methodological framework that jointly considers both patterns (block—or compartmental—structure and within-block nestedness). Our methodology can unveil the existence of IBN structures, as shown in Section III for a suitable benchmark. We then investigate, in Section IV, the question of how general (or anecdotal) such property is. We show that a large number of real datasets exhibit in-block nested structures that would have gone undetected under conventional modularity optimization/nestedness detection techniques. Our findings indicate that these previously-overlooked structures are in fact common in ecological and social systems. This opens a new direction for the structural analysis of ecological and social systems, discussed in Section V calling as well for new models to explain how IBN structures emerge.

II. DEFINITION AND QUANTIFICATION OF IN-BLOCK NESTEDNESS

In this Section, we develop a proper formulation of the problem of determining to what extent a given network is organized as loosely interconnected blocks, each of them internally nested. We begin by defining in a congruent manner both, global nestedness and the new in-block nestedness fitness \( \mathcal{I} \). In particular, maximizing the IBN fitness function \( \mathcal{I} \) allows us to unveil the best node partition in terms of IBN structure. We analyze synthetic networks to show that \( \mathcal{I} \)-maximization allows us to reconstruct ground truth IBN structures that would have gone undetected under the widely-used modularity optimization. We then proceed to analyze a large set of real networks – originated in the most varied disciplines – to evince that this type of structures is indeed a common occurrence in both uni- and bipartite networks of diverse nature.

A. A nestedness measure

Without loss of generality, consider a bipartite network, describing a relationship between two sets \( G = \{s, t, \ldots \} \) and \( \Gamma = \{\sigma, \tau, \ldots \} \) with cardinalities \( N_r \) and \( N_c \) respectively. The bipartite network can be represented as a binary adjacency matrix \( A \) whose elements are \( A_{s,t} = 1 \) if a relationship between elements \( s \) and \( t \) exists, or zero otherwise. In the same spirit as the measure Node Overlap-Decreasing Fill [27] (NODF), we introduce the global nestedness fitness \( \tilde{\mathcal{N}} \) which measures the amount of global overlap between row and column pairs

\[
\tilde{\mathcal{N}} = \frac{2}{N_r + N_c} \left\{ \sum_{s,t} \frac{O_{s,t}}{k_t (N_r - 1)} \Theta(k_s - k_t) \right. \\
+ \left. \sum_{\sigma,\tau} \frac{O_{\sigma,\tau}}{k_\tau (N_c - 1)} \Theta(k_\sigma - k_\tau) \right\},
\] (1)

\[
\tilde{\mathcal{N}} = \frac{2}{N_r + N_c} \left\{ \sum_{s,t} \frac{O_{s,t}}{k_t (N_r - 1)} \right. \\
+ \left. \sum_{\sigma,\tau} \frac{O_{\sigma,\tau}}{k_\tau (N_c - 1)} \right\},
\]
FIG. 2. Behavior of nestedness metrics $\tilde{N}$ and $N$, and modularity $Q$ in noiseless IBN synthetic networks. Panel (A) illustrates how global nestedness $\tilde{N}$ decreases with the number of blocks in a network—regardless of their internal perfectly nested structure. Notice that the shape parameter $\xi$ also affects $\tilde{N}$, as a side-effect of the strict decreasing condition imposed in the original definition of NODF. Panel (B) illustrates how modularity $Q$ increases with the number of blocks in a network. Panel (C) illustrates the kind of networks generated by the model (see Section III A). The vertical axis represents the shape parameter $\xi$, that quantifies the slimness of the nested structures. The horizontal axis stands for the number of modules. In non-integer settings (e.g. 2.5) there are an integer amount of nested blocks of the same size (e.g. 2) and one block of relative size equal to the remaining fraction (e.g. 0.5 of the size of the other communities). Panel (D) shows the global nestedness fitness $N$ which incorporates a null model, as defined by Eq. 5. Panel (E) shows the normalized variation of information between the modules detected by modularity optimization and the ground-truth (GT) blocks. Panel (F) shows the difference $M - B$ between the number $M$ of modules detected by the modularity optimization and the number $B$ of ground-truth blocks.

where $k_i$ corresponds to the degree of the element $i$ (regardless on whether it belongs to $G$ or $\Gamma$); $\Theta(\cdot)$ is the Heaviside step function (such that the only contributing terms are those in which the outer index has larger degree than the inner); $O_{st}$ measures the degree of overlap between row and column pairs as:

$$O_{s,t} = \sum_{v=1}^{N_r} A_{sv} A_{tv}, \quad O_{\sigma,\tau} = \sum_{u=1}^{N_c} A_{u\sigma} A_{u\tau}. \quad (2)$$

It is important to remark that Eq. 1 weighs linearly the contribution of rows and columns to $\tilde{N}$ (instead of quadratic weights as in NODF [27]). This is preferable when the difference between the number of rows and columns is considerable.

B. Recasting nestedness at the mesoscale level

We now introduce two new elements that allow us to generalize $N$ to the case where nested structures exist at a mesoscopic scale: a membership variable and a null model. First, we consider that both sets of nodes are partitioned into $C$ disjoint subsets, termed blocks. This implies that for each node $i$, it is possible to define a membership variable $\alpha_i$. Based on this, the total size of block $\ell$ can be obtained as

$$C(\ell) = \sum_{s=1}^{N_r} \delta(\alpha_s, \ell) + \sum_{\sigma=1}^{N_c} \delta(\alpha_\sigma, \ell), \quad (3)$$

where $\delta$ is the Kronecker Delta. In addition, $C_s = \sum_\sigma \delta(\alpha_s, \alpha_\sigma)$ and $C_\sigma = \sum_\tau \delta(\alpha_\sigma, \alpha_\tau)$ give, respectively, the number of nodes in the block nodes $s$ and $\sigma$ belong
to. The block overlap $O_{\cdot \cdot}$, now including the membership variable, can be obtained as

$$O_{s,t} = \sum_{u=1}^{N_c} A_{su} A_{tu} \delta(\alpha_s, \alpha_t),$$

$$O_{\sigma,\tau} = \sum_{u=1}^{N_r} A_{u\sigma} A_{u\tau} \delta(\alpha_u, \alpha_\sigma).$$

The null model gauges the expected overlap between a pair of nodes belonging to a class and aims at compensating the nestedness that can be explained solely by the nodes’ degrees. The expected overlap, given two nodes $s$ and $t$ with degrees $k_s$ and $k_t$, is obtained considering that the neighbors of each node are chosen uniformly at random. In this situation, the probability that both nodes share a common neighbor is simply $(k_s k_t)/N_c^2$ and therefore, the expected amount of shared neighbors, i.e. the expected overlap, is given by $\langle O_{s,t} \rangle = k_s k_t/N_c$. The same argument shows that for the columns, $\langle O_{\sigma,\tau} \rangle = k_{\sigma} k_{\tau}/N_r$.

We can now introduce the in-block nestedness fitness $\mathcal{I}$, which quantifies to which extent a network exhibits IBN,

$$\mathcal{I} = \frac{2}{N_r + N_c} \left\{ \sum_{s,t} \frac{O_{s,t} - \langle O_{s,t} \rangle}{k_t (C_s - 1)} \Theta(k_s - k_t) \delta(\alpha_s, \alpha_t) \right. + \left. \sum_{\sigma,\tau} \frac{O_{\sigma,\tau} - \langle O_{\sigma,\tau} \rangle}{k_{\tau} (C_{\tau} - 1)} \Theta(k_{\sigma} - k_{\tau}) \delta(\alpha_{\sigma}, \alpha_{\tau}) \right\}.$$

In this expression, some normalization factors have disappeared after straightforward simplifications. In the same spirit as in the row and column weighting of $\mathcal{N}$ (Eq. [1]), the per-block nestedness aggregates are weighted by the size of the block (i.e. $C_s$ and $C_\tau$). Notice that, for each pair of row nodes, $O_{s,t}$ only accounts for column nodes within the same block, while $\langle O_{s,t} \rangle$ considers all column nodes regardless of the block they belong to. This implies that, for any pair of rows, $\mathcal{I}$ will be in principle larger when they are assigned to the same block: in this case the difference $O_{s,t} - \langle O_{s,t} \rangle$ has positive contributions. On the other hand, the membership variable $\alpha$ allows to discard some of the comparisons, assigning row nodes to different communities. In general, an algorithm that correctly maximizes $\mathcal{I}$ will attempt to discard pairs whose contribution is negative to the aggregate. This intuition is equivalent for columns. The balance of such “mergesplit” strategy for rows and columns allows an algorithm to identify in-block nested structures by maximizing the objective function $\mathcal{I}$. Equation [3] is equally valid for multipartite networks, simply imposing that the two sets of nodes are identical. In this work, we have adopted a biologically-inspired optimization algorithm [29]. However, $\mathcal{I}$’s formulation—which closely follows that of modularity $Q$—enables the adoption of many existing heuristics (see [10] for an extensive review).

Noteworthy, the objective function $\mathcal{I}$ reduces to $\mathcal{N}$, corrected by a suitable null model, if one considers a single block ($\alpha_s = \alpha_t = \alpha$, $\forall s, \tau$), i.e.

$$\mathcal{N} = \frac{2}{N_r + N_c} \left\{ \sum_{s,t} \frac{O_{s,t} - \langle O_{s,t} \rangle}{k_t (N - 1)} \Theta(k_s - k_t) \right. + \left. \sum_{\sigma,\tau} \frac{O_{\sigma,\tau} - \langle O_{\sigma,\tau} \rangle}{k_{\tau} (M - 1)} \Theta(k_{\sigma} - k_{\tau}) \right\}.$$

While the difference between NODF and $\mathcal{N}$ is slight —except when $M \ll N$, or viceversa—, the null-model correction in $\mathcal{N}$ heavily alters the nestedness measure. In particular, note that fully connected nodes do not contribute to $\mathcal{N}$—as opposed to maximum contribution in the original formulation.

To illustrate this point, it is instructive to consider the case of one single nested block (Num. Blocks = 1 in Fig. 2A,C,D) (Note that Fig. 2 is explained in detail in the next Section). In this cut, the contrast between non-corrected measure of nestedness (panel 2A) and corrected (panel 2D) is very clear: when the perfectly nested network is very dense (bottom-left corner, $\xi < 2.5$), most of the nodes have large expected overlap with the few hubs. Hence, even though the nestedness condition is respected for all the pairs of nodes, the nestedness metric $\mathcal{N}$ only deviates little from its expected value under the null model, resulting in a small value of $\mathcal{N}$. In other words, the observed level of nestedness $\mathcal{N}$ can be simply explained by the network degree distribution. On the contrary, $\mathcal{N}$ and $\mathcal{N}$ are practically identical for $\xi > 2.5$. In this region, the slimmness of the nested structure and the strict decreasing connectivity condition (Heaviside function in Eqs. 1 and 4) heavily limit the value of nestedness, regardless the consideration of a null model ($\mathcal{N}$) or not ($\mathcal{N}$).

III. DETECTION OF IN-BLOCK NESTED STRUCTURES IN SYNTHETIC NETWORKS

In this Section, we first introduce a benchmark graph model with planted in-block nested structures (Section III A), and then present the results for the $Q$ and $\mathcal{I}$ optimization algorithms’ performance in reconstructing such planted structures (Section III B).
A. In-block nested structures generator

In a perfectly nested structure, rows (and columns) interact with a subset of the neighbors of the rows (and columns) of larger degree. Correctly ordering its adjacency matrix by row and column degree, it resembles an upper (possibly with some curvature) secondary diagonal matrix. Inspired by the p-norm unit ball equation, we synthetically generate such structures using

$$y = f_{\alpha}(x) = 1 - (1 - x^{1/\xi})^\xi$$  \hspace{1cm} (5)

where \(x \in [0,1]\) and \(\xi \in [1,\infty]\) dictates how stylized is the shape of the nested structure. The adjacency matrix of the nested structure with \(N_c\) nodes is constructed tessellating the \([0,1] \times [0,1]\) space into \(N_c^2\) squares; and then, adding a link into each matrix position whose center lies above the curve in Eq. (5) While an approach based on threshold graphs would have been also used, the degree sequences it produces are stochastic, introducing unnecessary fluctuations to the network generation [30].

Such noiseless nested structures are rarely found in real systems. Thus, on top of the previous scheme, we mimic random and uncorrelated noise using a dual-step procedure. In the first stage, we randomly remove links from the perfectly nested structure with probability \(p\). Given a network with \(E_c\) edges, \(pE_c\) of them will be removed in average. In the second step, the removed edges are randomly distributed across the empty elements of the adjacency matrix. These include initially empty positions (i.e. those lying below the function in Eq. (5) and empty positions resulting from the stochastic removal procedure. Parameter value \(p = 0\) corresponds to the initial structure and \(p = 1\) corresponds to an Erdős-Rényi network with average degree \(E_c/N_c\).

The construction of an adjacency matrix of an ideal in-block nested structure starts off with \(B\) (a real-valued number) blocks. Specifically, we build \([B]\) blocks of size \([N_c/B]\) and another with the remaining \(N_c - \lfloor N_c/B \rfloor\) nodes. In the previous, \(\lfloor \cdot \rfloor\) stands for the integer part function. In this way, the network produced has some level of heterogeneity (albeit the size of all blocks remains in the same order of magnitude). Then \(A_T\), can be constructed repeating the procedure for each block, \(A^c\), and joining them to compose a block diagonal matrix

$$A_T = \begin{bmatrix}
A^{c_1} & A^{c_{11}} & \cdots & A^{c_{1B}} \\
A^{c_2} & A^{c_{21}} & \cdots & A^{c_{2B}} \\
\vdots & \vdots & \ddots & \vdots \\
A^{c_{11}} & A^{c_{21}} & \cdots & A^{c_{BB}}
\end{bmatrix},$$  \hspace{1cm} (6)

where \(A^o = 0\) is a matrix of the required size.

Similarly to the intra-block noise, we reproduce inter-block perturbations with an additional dual-step procedure controlled by the parameter \(\mu \in [0,1]\). In this first step, for each block, each link is removed with probability \(p_i = \mu (B - 1)/B\). In the second step, those links are distributed at random to connect one node of the original block with a random node of a different block. Probability \(p_i\) depends on the number of blocks, since our purpose is that for \(\mu = 1\) the amount of links within each block is the same as the amount of links connecting any two distinct blocks. In the limit situation of \(p = \mu = 0\) the outcome corresponds to a noiseless in-block nested structure and for \(p = \mu = 1\) the outcome corresponds to an Erdős-Rényi network with the same average degree as the germinal noiseless structure, see Appendix [A].

The described generative process can be locally implemented in terms of edge probabilities. Along these lines, the probability of having a link between nodes \(i\) and \(j\) within a block becomes

$$P(A^c_{ij}) = (1 - p + pp_r) \Theta (N_c - f_o(i N_o)) + (1 - p_r) (1 - \Theta (N_c - f_o(i N_o))) (1 - p_i),$$  \hspace{1cm} (7)

where \(\Theta\) is the Heaviside function. The term within square brackets is related to the intra-block noise. In the first term, \((1 - p)\) corresponds to the probability of not altering the link. The second, \(pp_r\), corresponds to the probability of recovering a link, after removal, in the random dispersion of removed links. These two terms are restricted, by \(\Theta\) function, to the region where links exist in the noiseless structure. The third term, \(p_r = pE_c (N_c - E_c + pE_c)^{-1}\), corresponds to the probability of selecting link \(A^c_{ij}\) in the random distribution of removed links. Eventually, the term \((1 - p_i)\) corresponds to the probability of not removing the link in the process of generating inter-block noise.

The probability of a inter-block link is

$$P(A^o_{ij}) = \frac{2E_c p_i}{2(B - 1)N_c^2} = \frac{\mu E_c}{N_c^2 B}.$$  \hspace{1cm} (8)

The numerator accounts for the amount of removed links from the blocks related to the off-diagonal block \(A^o\), that is compartment \(i\) and \(j\). The denominator accounts for the possible places where each of those links can be placed. Note that the 2 on both numerator and denominator explicitly shows that each removed link of compartment \(A^c\) can be reallocated in \(A^{oc}\) or \(A^{oc}\).

The noisy version given by Eqs. (7) and (8) of the original noiseless in-block nested structure generates a network with equivalent average degree. This is formally proved in Appendix [A]. An example of these synthetically generated structures is shown in Fig. 1 and Fig. 3.

B. IBN optimization applied to synthetic networks

In Fig. 2, we unveil the limitations of current techniques to detect IBN structures in noiseless networks (IBN networks) where no links exist between nodes belonging to different blocks. In a more realistic setting, where links can connect nodes that belong to different blocks, such weaknesses become even more apparent (see Fig. 5).

We generated synthetic (unipartite) networks of \(N_c\) nodes and \(E_c\) edges within blocks, where the level of in-block nestedness and the number of inter-block links can
be varied in a controlled manner by means of few parameters (see Methods): the number of blocks $B$, the shape parameter $\xi$, the in-block nestedness parameter $p$, and the mixing parameter $\mu$. To allow for heterogeneity in block size, we build $\lfloor B \rfloor$ blocks of size $\lfloor N_c/B \rfloor$ and another with the remaining nodes. $\xi$ determines the density of the network, controlling how stylized the nested structure is. The in-block nestedness parameter $p$ gives the fraction of links that do not respect the notion of perfectly nested organization within a block and the mixing parameter $\mu$ measures the fraction of inter-block links.

The experiments in Fig. 2 correspond to $3 \times 10^4$ networks generated with parameters $p = \mu = 0$, with varying number of blocks and number of edges (see Fig. 2C for an illustration of the resulting adjacency matrices). By construction all networks have maximum $I$, in the range $(0.16, 0.90)$, depending on the shape parameter $\xi$ and number of blocks $B$. Unsurprisingly, modularity $Q$ increases as the number of block increases (Fig. 2B).

We find that in this setting, the modules detected by the modularity-maximization algorithm can be very different from the planted blocks, as measured by Normalized Variation of Information (NVI), Fig. 2E. The difference between detected and planted blocks is larger for sparser networks (upper region of Fig. 2E). Fig. 2F shows the difference between the number $M$ of detected modules by the $Q$-maximization and the number $B$ of planted blocks. The difference is non-zero for a large region of the parameter space. In particular, the modularity optimization algorithm detects more than one module in a network composed of one single block with internal nested structure (see Fig. 2F, left-corner). This happens because the modularity-optimization algorithm tends to form a module that only contains the nodes with largest degree. Figs. 2E-F indicate that modularity optimization is only reliable in the limit of large number of blocks and dense networks (lower-right corner).

The results of Figs. 2B,E,F make clear that measuring
modularity and nestedness as two independent network properties is inherently flawed: modularity-optimization algorithms detect more than one module in a network composed of a single nested block, and the NVI between detected modules and planted blocks is in general large. We have verified that the $I$-optimization algorithm introduced in this paper overcomes these limitations and is able to correctly recover the planted structure for all the parameter values shown in Fig. 2.

Figure 3 shows the results for an exhaustive exploration of the $(p, \mu)$ parameter space over 2600 networks with a fixed $B = 3$ and shape parameter $\xi = 4$ (see Fig. 3A for an illustration of the resulting adjacency matrices). Results on these synthetic networks after a modularity optimization process (Fig. 3B) show that $Q$ is almost insensitive to changes in the parameters of the model: notice that its range is quite narrow, $0.55 < Q < 0.7$, and only mildly affected by the $p$ parameter (i.e., by the level of IBN). This itself is a consequence of the fact that $Q$ does not consider any particular structure within the blocks, but only their internal density. Even further, the value of modularity fluctuates around $Q \approx 0.6$, in remarkable accordance with the predictions in Guimerà et al. [33], for sparse graphs, like those obtained with $\xi = 4$.

Figure 3A shows the value of in-block nestedness fitness, $I$, after a maximization procedure based on a biologically-inspired optimization algorithm (see Methods). Evidently, $I$ is sensitive to both the modular structure, and the nested organization within, taking a maximum value for $\mu = 0$ and $p = 0$. When we increase the randomness in either dimension the obtained in-block nestedness fitness smoothly decreases, reaching a global minimum when $\mu = 1$ and $p = 1$.

Since it can be assumed that the block structure is known a priori, it is possible to quantify how far a given partition is from the planted one. As before, we resort on NVI to assess the quality of the partitions obtained optimizing $I$ (Eq. 1) and the quality of the partitions obtained maximizing $Q$ [31]. Focusing on the quality of the Q-detected partition with respect to the prescribed one (Fig. 3C), we emphasize that modularity does not recover the planted partition in any parameter configuration, not even at $\mu = 0$. Remarkably, changes in NVI are independent of the parameter $p$, related to the level of disorder within each block. In contrast, we see that $I$ optimization allows to unveil the planted partition for a region along the $\mu$ axis, as long as $p$ remains low (Fig. 3B): the presence of internal nestedness compensates the tenuous identity of the blocks, caused by large $\mu$. The parameter region corresponding to low $p$ and large $\mu$ is also the region where Q-detected partitions and $I$-detected partitions differ the most (Fig. 3D). This points out that the Q-detected partitions are particularly unreliable when there is a clear internal nested structure and there exist a significant number of inter-block links.

IV. DETECTION OF IN-BLOCK NESTED STRUCTURES IN REAL DATASETS

The previous sections demonstrate the adequacy and robustness of $I$ –and the inherent flaws of modularity $Q$ and global nestedness $N$– to unveil IBN structures. However, those analyses would be limited to a mere academic exercise in absence of ample (in terms of examples and origin) empirical evidence. To demonstrate the practical aspects of the proposed methodology, we have analyzed a total of 334 networks, including both unipartite (57) and bipartite (277) ones which are known to display some level of nested organization. Most of them (209 bipartite networks) belong to ecology [34]–mostly mutualistic networks– and the rest belong to online platforms (68 bipartite networks) and social networks (57 unipartite networks). Table 1 in Appendix B details the origin and characteristics of each dataset.

As a visual intuition, Fig. 4 displays the adjacency matrix of four of these networks, where rows and columns have been sorted following different criteria: for left and central columns ($I$- and $Q$-maximizing partitions, respectively), nodes in the same block are placed together, and they are ranked by degree (within blocks) to make more apparent a possible IBN structure; in the right column, nodes are simply ranked by degree. Panel A shows such arrangements for a host-parasite network (see A_HP_050 in Table 1 of the Appendix B). Clearly, the three matrix representations look very different. In this case, $I$ favors the existence of a large, highly nested block, and a set of smaller clusters with a clear internal organization as well, whereas $Q$ renders several, similarly sized, highly dense modules with no clear internal nested organization. Even though the classical NODF measure hints at some degree of global nestedness, taking into account the null model ($N = 0.059$) seems to indicate that the nested organization is a simple consequence of the network’s degree distribution. Panel B shows the results for a pollination mutualistic network (see M_PL_001 in Table 1 of the Appendix B). The system exhibits a clear IBN structure that cannot be detected through the maximization of modularity. From the results in panels A and B, it is worth remarking that the observation of IBN structures in ecosystems with different types of interactions demands a reconsideration of which patterns should or should not be expected in them. Panel C shows the results for a urban user-service network (see Chennai in Table 1 of the Appendix B). We observe again that global nestedness fails to characterize the predominant organization of the system, i.e. an IBN structure. Panel D shows the results for a unipartite network representing friendship relations in a Dutch school class (see c2 in Table 1 of the Appendix B). The conclusions of the analysis of this network are similar to the ones in Panel A.

While Fig. 4 is helpful to get an intuition of how $I$, $Q$ and $N$ work, these representations may be misleading. For example, the degree-ranked representation in Panel B (right) conveys the idea that this particular plant-
FIG. 4. Interaction matrices of four illustrative systems: (A) a host-parasite competitive bipartite network in Volga-Kama Nature Reserve; (B) a pollination mutualistic bipartite network in Cordón del Cepo, Chile; (C) a urban bipartite network accounting for citizen visits to city services in Chennai, India; and (D) a unipartite friendship network in a Dutch school class. For all of them, rows and columns have been arranged to highlight different aspects: block membership and degree for $I$-(left) and $Q$-maximization (center) partitions; global degree ranking in the right column.

pollinator network is clearly (and globally) nested –but the other two arrangements are qualitatively convincing as well. For this reason, we have systematically compared the results of $N$ and $Q$, on one side, and $I$, on the other, for the whole set of real networks mentioned above. In Fig. 5A, two color-coded scatter plots are shown for uni-(left) and bipartite (right) networks. Strikingly, modularity $Q$ and in-block nestedness $I$ are not strongly correlated in real datasets. Networks that exhibit small or intermediate values of modularity (compatible with those of a random network [33]) may show high $I$, regardless of the $N$ score. Also, large values of $Q$ –which unsurprisingly display nestedness $\approx 0$– indeed exhibit both large and small $I$ scores as well. Beyond the uncorrelated behavior between the three descriptors, what surfaces here is the fact that when analyzing data we may be overlooking a relevant pattern –IBN structure– just because two partial views of it ($N$ and $Q$ taken independently) appear to be non-significant.

Further, in Fig. 5B we show the value of $I$ for partitions obtained by maximizing $Q$ confronted to the maximization of $I$ itself. This plot evidences that modularity optimization may sometimes render partitions which do have some in-block organization (near the diagonal), but most often it is blind to it. This result highlights that using an approach where modularity is maximized, to successively evaluate nestedness within the blocks identified (i.e. the approach in [22, 24]), is not able to unveil the IBN structure in most real-world networks.

V. CONCLUSIONS

The emergence of structural patterns in complex networks is a consequence of the dynamics that take place on them. While the ultimate goal is to understand how these dynamics operate, this is not feasible until the correct methods to identify those patterns are available. The increasing evidence that nestedness and modularity appear in many empirical contexts; the already abundant hints that they may appear together; and the importance of both to disentangle how they affect –and are affected by– the evolution of a system, prompt the need of rethinking the strategy to detect the occurrence of in-block nestedness. Inspired by the NODF and modularity optimization, in this paper we have developed a methodology to detect in-block nested structures. The objective function at the core of this method naturally embeds a suitable null model to discount the in-block nestedness of the network that can be ascribed to randomness. Beyond the formal correctness of our formulation, demonstrated by means of a suitable benchmark to generate synthetic networks, we have shown that it overcomes the inherent limitations of nestedness and modularity (as independent methods) for this task. In structural terms, our approach can be interpreted as a generalization of the concept of nestedness (as expressed in NODF), re-framing it to the mesoscopic network scale. Along this line, the door remains open to further development. For example, slight modifications on our formulation generalize the notion of multiple core-periphery structures, which has been recently addressed from a different starting point [36]. Other directions may be related to identification techniques (e.g. stochastic block models [37]) or the design of distinct, specialized null models [38, 39].

The analysis of real data has shown that many networks display in-block nested structure, regardless of their NODF and modularity scores. This finding suggests that previous works may have overlooked important features when discussing the organization of real systems. The existence of in-block nested structures affects the debate around population dynamics, in ecology specially,
FIG. 5. (A) Scatter plot confronting modularity and nestedness with the I measure (color coded). Each point represents either an ecological, urban or social network (left panel: unipartite networks; right panel: bipartite networks). Note that bipartite networks have been analyzed under the formulation of Barber’s modularity [35]. (B) Comparison of in-block nestedness value obtained optimising modularity or directly I, as we propose.

in terms of which patterns maximize survival [5], and why. Our methodological contribution thus uncovers the need for models – beyond host-parasite [40, 41] – that explain how networks transition between possible configurations: from modular, to combined, to purely nested architectures, as suggested by the emergence of collective attention processes [9]; or from nested to combined architectures, as one would expect in a growing, but highly structured, system with increasing specialization. Note-worthy, it is not clear whether these processes are reversible, as they may respond to different system-wide adaptive processes.

ACKNOWLEDGMENTS

CJT and MSM acknowledges the University of Zürich for support through the URPP Social Networks. MSM also acknowledges financial support from the Swiss National Science Foundation Grant No. 200020-156188.

Appendix A: Average degree of the synthetic network generation model

The synthetic in-block nested network generator we propose in the main text include two parameters to generate noise. Only as a summary, parameter p controls the intra-compartment noise and parameter µ controls the inter-community noise. The noisy versions of the initial germinal noiseless in-block nested structure has the same average degree. We first show that adding intra-community noise does not alter the average number of edges. Under this situation µ = 0 and Eq. 7 becomes

\[
P(A_{ij}^c) = ((1 - p) + pp_r)\Theta(jN_c - f_n(iN_c)) + p_r(1 - \Theta(jN_c - f_n(iN_c)))\tag{A1}
\]

and the expected number of edges within a compartment is \(\langle E_c \rangle = \sum_{ij} P(A_{ij}^c)\). Reorganizing common terms of the sums we obtain that

\[
\langle E_c \rangle = \sum_{ij} P(A_{ij}) =

= ((1 - p) + pp_r)\sum_{ij} \Theta(jN_c - f_n(iN_c)) \tag{A2}

+ p_r \sum_{ij} (1 - \Theta(jN_c - f_n(iN_c)))
\]

The heaviside step function is one when a link was existing in the original noiseless compartment. Thus, by
definition, \( E_c = \sum_{ij}^{N_c} \Theta(jN_c - f_n(iN_c)) \), and consequently
\[ N_c - E_c = \sum_{ij}^{N_c} (1 - \Theta(jN_c - f_n(iN_c))). \]
Substituting, this to Eq. A2 leads to
\[
\left< E_c \right> = (1 - p + pp_r) E_c + p_r(N_c^2 - E_c) =
E_c - pE + p_r(N_c - E_c + pE_c) = E_c
\] (A3)

We now consider the case where \( \mu \neq 0 \). The expected amount of edges within the full in-block nested structure
\[
\left< E \right> = B \sum_{ij}^{N_c} P(A_{ij}) +
B(B - 1) \sum_{ij}^{N_c} P(A_{ij}^*) =
B(1 - p_i) \sum_{ij}^{N_c} [(1 - p + pp_r) \Theta(jN_c - f_n(iN_c))] +
B(1 - p_i) \sum_{ij}^{N_c} [p_r \Theta(f_n(iN_c) - jN_c)]
E_c \mu
\] (A4)

Q.E.D.

Appendix B: Real datasets used in the experiments

The set of real networks used throughout the article comprise ecological and social systems. The largest subset – ecological networks [34] – represent mutualistic and competitive systems, including macroscopic and microscopic environments. The analyzed social systems include social communication networks such as face-to-face interactions, e-mail contacts and Twitter messages; urban systems such as user check-ins to city services (museum, market, restaurant, etc.); technological systems such as cooperative software development projects, where we account which files each user works on. Some of these networks have been previously shown to exhibit nestedness and modularity jointly [9, 24, 25]. However, none of them have not been previously analyzed in the proposed setting.

| Name   | M | N  | Connectance | Type           | Relation type               |
|--------|---|----|-------------|-----------------|-----------------------------|
| A_HP_002 | 24 | 18 | 0.22        | bipartite       | Ecological Host-Parasite    |
| A_HP_003 | 9  | 23 | 0.52        | bipartite       | Ecological Host-Parasite    |
| A_HP_006 | 37 | 16 | 0.21        | bipartite       | Ecological Host-Parasite    |
| A_HP_008 | 24 | 8  | 0.19        | bipartite       | Ecological Host-Parasite    |
| A_HP_009 | 22 | 14 | 0.31        | bipartite       | Ecological Host-Parasite    |
| A_HP_010 | 31 | 18 | 0.16        | bipartite       | Ecological Host-Parasite    |
| A_HP_013 | 22 | 11 | 0.43        | bipartite       | Ecological Host-Parasite    |
| A_HP_018 | 21 | 15 | 0.40        | bipartite       | Ecological Host-Parasite    |
| A_HP_020 | 17 | 16 | 0.39        | bipartite       | Ecological Host-Parasite    |
| A_HP_022 | 18 | 16 | 0.24        | bipartite       | Ecological Host-Parasite    |
| A_HP_025 | 40 | 18 | 0.15        | bipartite       | Ecological Host-Parasite    |
| A_HP_026 | 18 | 15 | 0.53        | bipartite       | Ecological Host-Parasite    |
| A_HP_027 | 30 17 0.21 bipartite | Ecological Host-Parasite |
| A_HP_029 | 34 15 0.15 bipartite | Ecological Host-Parasite |
| A_HP_031 | 31 25 0.22 bipartite | Ecological Host-Parasite |
| A_HP_033 | 25 22 0.36 bipartite | Ecological Host-Parasite |
| A_HP_037 | 21 17 0.25 bipartite | Ecological Host-Parasite |
| A_HP_042 | 32 21 0.12 bipartite | Ecological Host-Parasite |
| A_HP_043 | 29 9 0.28 bipartite | Ecological Host-Parasite |
| A_HP_044 | 26 27 0.28 bipartite | Ecological Host-Parasite |
| A_HP_046 | 39 17 0.30 bipartite | Ecological Host-Parasite |
| A_HP_047 | 26 11 0.35 bipartite | Ecological Host-Parasite |
| A_HP_050 | 35 27 0.24 bipartite | Ecological Host-Parasite |
| A_HP_051 | 26 13 0.32 bipartite | Ecological Host-Parasite |
| A_PH_004 | 22 52 0.16 bipartite | Ecological Plant-Herbivore |
| A_PH_005 | 24 54 0.13 bipartite | Ecological Plant-Herbivore |
| A_PH_006 | 88 6 0.22 bipartite | Ecological Plant-Herbivore |
| A_PH_007 | 64 5 0.30 bipartite | Ecological Plant-Herbivore |
| M_PA_003 | 15 24 0.12 bipartite | Ecological Plant-Ant |
| M_PA_004 | 48 41 0.14 bipartite | Ecological Plant-Ant |
| M_PL_001 | 101 84 0.04 bipartite | Ecological Pollination |
| M_PL_002 | 64 43 0.07 bipartite | Ecological Pollination |
| M_PL_003 | 25 36 0.09 bipartite | Ecological Pollination |
| M_PL_004 | 102 12 0.14 bipartite | Ecological Pollination |
| M_PL_005 | 275 96 0.03 bipartite | Ecological Pollination |
| M_PL_006 | 61 17 0.14 bipartite | Ecological Pollination |
| M_PL_007 | 36 16 0.15 bipartite | Ecological Pollination |
| M_PL_008 | 38 11 0.25 bipartite | Ecological Pollination |
| M_PL_009 | 118 24 0.09 bipartite | Ecological Pollination |
| M_PL_010 | 76 31 0.19 bipartite | Ecological Pollination |
| M_PL_012 | 55 29 0.09 bipartite | Ecological Pollination |
| M_PL_013 | 56 9 0.20 bipartite | Ecological Pollination |
| M_PL_014 | 81 29 0.08 bipartite | Ecological Pollination |
| M_PL_015 | 666 131 0.03 bipartite | Ecological Pollination |
| M_PL_016 | 179 26 0.09 bipartite | Ecological Pollination |
| M_PL_017 | 79 25 0.15 bipartite | Ecological Pollination |
| M_PL_018 | 105 39 0.09 bipartite | Ecological Pollination |
| M_PL_019 | 85 40 0.08 bipartite | Ecological Pollination |
| M_PL_020 | 91 20 0.10 bipartite | Ecological Pollination |
| M_PL_021 | 677 91 0.02 bipartite | Ecological Pollination |
| M_PL_022 | 45 21 0.09 bipartite | Ecological Pollination |
| M_PL_023 | 72 23 0.08 bipartite | Ecological Pollination |
| M_PL_025 | 44 13 0.25 bipartite | Ecological Pollination |
| M_PL_026 | 54 105 0.04 bipartite | Ecological Pollination |
| M_PL_027 | 60 18 0.11 bipartite | Ecological Pollination |
| M_PL_028 | 139 41 0.07 bipartite | Ecological Pollination |
| M_PL_029 | 118 49 0.06 bipartite | Ecological Pollination |
| M_PL_030 | 53 28 0.07 bipartite | Ecological Pollination |
| M_PL_031 | 49 48 0.07 bipartite | Ecological Pollination |
| M_PL_032 | 33 7 0.28 bipartite | Ecological Pollination |
| M_PL_033 | 34 13 0.32 bipartite | Ecological Pollination |
| M_PL_034 | 128 26 0.09 bipartite | Ecological Pollination |
| M_PL_035 | 36 61 0.08 bipartite | Ecological Pollination |
| M_PL_037 | 40 10 0.18 bipartite | Ecological Pollination |
| M_PL_038 | 42 8 0.24 bipartite | Ecological Pollination |
| M_PL_039 | 17 18 0.15 bipartite | Ecological Pollination |
| M_PL_040 | 43 29 0.09 bipartite | Ecological Pollination |
| M_PL_041 | 43 31 0.11 bipartite | Ecological Pollination |
| M_PL_043 | 82 28 0.11 bipartite | Ecological Pollination |
| M_PL_044 | 609 110 0.02 bipartite | Ecological Pollination |
| M_PL_045 | 26 17 0.14 bipartite | Ecological Pollination |
| M_PL_046 | 44 16 0.39 bipartite | Ecological Pollination |
| M_PL_047 | 186 19 0.12 bipartite | Ecological Pollination |
| M_PL_048 | 236 | 30 | 0.09 | bipartite | Ecological Pollination |
| M_PL_049 | 225 | 37 | 0.07 | bipartite | Ecological Pollination |
| M_PL_050 | 35 | 14 | 0.18 | bipartite | Ecological Pollination |
| M_PL_051 | 90 | 14 | 0.13 | bipartite | Ecological Pollination |
| M_PL_052 | 39 | 15 | 0.16 | bipartite | Ecological Pollination |
| M_PL_053 | 294 | 99 | 0.02 | bipartite | Ecological Pollination |
| M_PL_054 | 318 | 113 | 0.02 | bipartite | Ecological Pollination |
| M_PL_055 | 195 | 64 | 0.03 | bipartite | Ecological Pollination |
| M_PL_056 | 365 | 91 | 0.03 | bipartite | Ecological Pollination |
| M_PL_057 | 883 | 114 | 0.02 | bipartite | Ecological Pollination |
| M_PL_058 | 83 | 32 | 0.12 | bipartite | Ecological Pollination |
| M_PL_060_01 | 39 | 11 | 0.22 | bipartite | Ecological Pollination |
| M_PL_060_02 | 38 | 12 | 0.23 | bipartite | Ecological Pollination |
| M_PL_060_03 | 45 | 13 | 0.22 | bipartite | Ecological Pollination |
| M_PL_060_04 | 46 | 21 | 0.14 | bipartite | Ecological Pollination |
| M_PL_060_05 | 54 | 33 | 0.08 | bipartite | Ecological Pollination |
| M_PL_060_06 | 45 | 26 | 0.08 | bipartite | Ecological Pollination |
| M_PL_060_07 | 39 | 29 | 0.10 | bipartite | Ecological Pollination |
| M_PL_060_08 | 28 | 19 | 0.14 | bipartite | Ecological Pollination |
| M_PL_060_09 | 40 | 18 | 0.11 | bipartite | Ecological Pollination |
| M_PL_060_10 | 25 | 14 | 0.15 | bipartite | Ecological Pollination |
| M_PL_060_11 | 20 | 14 | 0.15 | bipartite | Ecological Pollination |
| M_PL_060_12 | 26 | 11 | 0.21 | bipartite | Ecological Pollination |
| M_PL_060_13 | 31 | 7 | 0.22 | bipartite | Ecological Pollination |
| M_PL_060_14 | 37 | 11 | 0.24 | bipartite | Ecological Pollination |
| M_PL_060_15 | 37 | 14 | 0.25 | bipartite | Ecological Pollination |
| M_PL_060_16 | 39 | 17 | 0.17 | bipartite | Ecological Pollination |
| M_PL_060_17 | 35 | 17 | 0.17 | bipartite | Ecological Pollination |
| M_PL_060_18 | 28 | 20 | 0.13 | bipartite | Ecological Pollination |
| M_PL_060_19 | 13 | 18 | 0.18 | bipartite | Ecological Pollination |
| M_PL_060_22 | 31 | 13 | 0.16 | bipartite | Ecological Pollination |
| M_PL_060_23 | 27 | 12 | 0.18 | bipartite | Ecological Pollination |
| M_PL_060_24 | 24 | 14 | 0.14 | bipartite | Ecological Pollination |
| M_PL_061_05 | 22 | 12 | 0.19 | bipartite | Ecological Pollination |
| M_PL_061_06 | 24 | 11 | 0.22 | bipartite | Ecological Pollination |
| M_PL_061_07 | 23 | 9 | 0.30 | bipartite | Ecological Pollination |
| M_PL_061_19 | 21 | 11 | 0.19 | bipartite | Ecological Pollination |
| M_PL_061_23 | 26 | 10 | 0.17 | bipartite | Ecological Pollination |
| M_PL_061_38 | 23 | 10 | 0.21 | bipartite | Ecological Pollination |
| M_PL_061_40 | 26 | 9 | 0.25 | bipartite | Ecological Pollination |
| M_PL_061_45 | 23 | 11 | 0.20 | bipartite | Ecological Pollination |
| M_PL_061_46 | 21 | 13 | 0.16 | bipartite | Ecological Pollination |
| M_PL_061_47 | 26 | 9 | 0.24 | bipartite | Ecological Pollination |
| M_PL_063 | 9 | 55 | 0.25 | bipartite | Ecological Pollination |
| M_SD_002 | 9 | 31 | 0.43 | bipartite | Ecological Seed Dispersal |
| M_SD_003 | 16 | 25 | 0.17 | bipartite | Ecological Seed Dispersal |
| M_SD_004 | 20 | 34 | 0.14 | bipartite | Ecological Seed Dispersal |
| M_SD_005 | 13 | 25 | 0.15 | bipartite | Ecological Seed Dispersal |
| M_SD_006 | 15 | 21 | 0.16 | bipartite | Ecological Seed Dispersal |
| M_SD_007 | 7 | 72 | 0.28 | bipartite | Ecological Seed Dispersal |
| M_SD_010 | 14 | 50 | 0.33 | bipartite | Ecological Seed Dispersal |
| M_SD_012 | 29 | 35 | 0.14 | bipartite | Ecological Seed Dispersal |
| M_SD_013 | 19 | 36 | 0.29 | bipartite | Ecological Seed Dispersal |
| M_SD_014 | 17 | 16 | 0.44 | bipartite | Ecological Seed Dispersal |
| M_SD_015 | 27 | 5 | 0.04 | bipartite | Ecological Seed Dispersal |
| M_SD_016 | 61 | 24 | 0.34 | bipartite | Ecological Seed Dispersal |
| M_SD_018 | 32 | 29 | 0.07 | bipartite | Ecological Seed Dispersal |
| M_SD_019 | 40 | 169 | 0.10 | bipartite | Ecological Seed Dispersal |
| M_SD_020 | 33 | 25 | 0.18 | bipartite | Ecological Seed Dispersal |
| M_SD_021 | 28 | 18 | 0.26 | bipartite | Ecological Seed Dispersal |
| M_SD_022 | 110 | 207 | 0.05 | bipartite | Ecological Seed Dispersal |
| Women Event Participation | 18 | 14 | 0.35 | bipartite | Social | 12 |
|---------------------------|----|----|------|----------|-------|----|
| Athens                    | 545| 224| 0.02 | bipartite| Foursquare Urban Checkins | 43 | 44 |
| Chennai                   | 280| 279| 0.04 | bipartite| Foursquare Urban Checkins | 43 | 44 |
| Mcp2000                   | 146| 1131| 0.13 | bipartite| Github Project |
| animatecss                | 77 | 254| 0.03 | bipartite| Github Project |
| html5-boilerplate        | 240| 382| 0.01 | bipartite| Github Project |
| javascript                | 285| 70| 0.03 | bipartite| Github Project |
| UsrHstg_0_1439            | 49 | 48| 0.08 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_10080_11519       | 174| 373| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_10800_12239       | 160| 317| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_11520_12959       | 177| 349| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_12240_13679       | 158| 332| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_12960_14399       | 147| 381| 0.02 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_13680_15119       | 170| 404| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_14240_15839       | 215| 418| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_14400_2879        | 148| 344| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_15120_16559       | 342| 955| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_15840_17279       | 378| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_16560_17999       | 283| 872| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_17280_18719       | 209| 428| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_18000_19439       | 192| 378| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_18720_20159       | 167| 429| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_19440_20879       | 253| 709| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_20160_21599       | 403| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_20880_22319       | 407| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_21600_23039       | 386| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_21600_35999       | 170| 398| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_22320_23759       | 441| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_23040_24479       | 418| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_23760_25199       | 354| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_24480_25919       | 314| 844| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_25200_26639       | 346| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_25920_27359       | 414| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_26640_28079       | 429| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_27360_28799       | 442| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_28080_29519       | 477| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_28800_30239       | 713| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_28800_4319        | 172| 365| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_29520_30959       | 909| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_30240_31679       | 870| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_30960_32399       | 960| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_31680_33119       | 1075| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_32400_33839       | 1115| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_33120_34559       | 1244| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_33840_35279       | 1420| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_34560_35999       | 1540| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_35280_36719       | 1480| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_36000_37439       | 1516| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_37440_38879       | 1635| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_38160_39599       | 1407| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_38880_40319       | 1347| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_39600_41039       | 1331| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_40320_41759       | 1180| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_41040_42479       | 1093| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_41760_43199       | 1059| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_42480_43919       | 1049| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_43200_44639       | 999| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_43920_45359       | 145| 273| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_44640_46079       | 820| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| UsrHstg_45360_46799       | 611| 1024| 0.01 | bipartite| Twitter User-Hashtag | 9 |
| Node Id        | Degree | Fraction | Type         | Label                          |
|---------------|--------|----------|--------------|-------------------------------|
| UsrHstg_5040_6479 | 149    | 0.01     | bipartite    | Twitter User-Hashtag          |
| UsrHstg_5700_7199 | 134    | 0.01     | bipartite    | Twitter User-Hashtag          |
| UsrHstg_6480_7919 | 116    | 0.02     | bipartite    | Twitter User-Hashtag          |
| UsrHstg_7200_8639 | 107    | 0.02     | bipartite    | Twitter User-Hashtag          |
| UsrHstg_720_2159 | 76     | 0.03     | bipartite    | Twitter User-Hashtag          |
| UsrHstg_7920_9359 | 107    | 0.02     | bipartite    | Twitter User-Hashtag          |
| UsrHstg_8640_10079 | 133    | 0.01     | bipartite    | Twitter User-Hashtag          |
| bunt0         | 31     | 0.01     | unipartite   | Social contacts               |
| bunt1         | 32     | 0.14     | unipartite   | Social contacts               |
| bunt3         | 32     | 0.21     | unipartite   | Social contacts               |
| bunt4         | 32     | 0.25     | unipartite   | Social contacts               |
| bunt5         | 32     | 0.29     | unipartite   | Social contacts               |
| bunt6         | 32     | 0.24     | unipartite   | Social contacts               |
| c1            | 26     | 0.19     | unipartite   | Social contacts               |
| c2            | 26     | 0.25     | unipartite   | Social contacts               |
| c3            | 26     | 0.31     | unipartite   | Social contacts               |
| c4            | 26     | 0.28     | unipartite   | Social contacts               |
| E-mail contacts (Milchaski) | 196   | 0.10     | unipartite   | Social contacts               |
| stu98t0       | 31     | 0.01     | unipartite   | Social contacts               |
| stu98t2       | 34     | 0.27     | unipartite   | Social contacts               |
| stu98t3       | 34     | 0.36     | unipartite   | Social contacts               |
| stu98t5       | 34     | 0.35     | unipartite   | Social contacts               |
| stu98t6       | 34     | 0.35     | unipartite   | Social contacts               |
| Primary School | 242    | 0.28     | unipartite   | Social contacts               |
| High School 2011 | 126   | 0.22     | unipartite   | Social contacts               |
| High School 2012 | 180   | 0.14     | unipartite   | Social contacts               |
| E-mail contacts | 1133  | 0.01     | unipartite   | Social                        |
| zachary       | 34     | 0.13     | unipartite   | Social                        |
| Enron E-mail contacts M12 | 139   | 0.02     | unipartite   | Social                        |
| Enron E-mail contacts M13 | 287   | 0.01     | unipartite   | Social                        |
| Enron E-mail contacts M14 | 383   | 0.01     | unipartite   | Social                        |
| Enron E-mail contacts M15 | 420   | 0.01     | unipartite   | Social                        |
| Enron E-mail contacts M16 | 286   | 0.01     | unipartite   | Social                        |
| Enron E-mail contacts M17 | 367   | 0.01     | unipartite   | Social                        |
| Enron E-mail contacts M18 | 418   | 0.01     | unipartite   | Social                        |
| Enron E-mail contacts M19 | 1052  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M20 | 1285  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M21 | 2485  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M22 | 2477  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M23 | 2081  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M24 | 2158  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M25 | 3106  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M26 | 3479  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M27 | 3491  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M28 | 3990  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M29 | 4291  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M30 | 4341  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M31 | 4344  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M32 | 4081  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M33 | 3810  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M34 | 4341  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M35 | 7287  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M36 | 5013  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M37 | 5013  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M38 | 4384  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M39 | 4702  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M40 | 7335  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M41 | 2974  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M42 | 3109  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M43 | 2231  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M44 | 1639  | 0.00     | unipartite   | Social                        |
| Enron E-mail contacts M45 | 313   | 0.01     | unipartite   | Social                        |
An idea based on honey bee swarm for numerical optimization, unipartite
et al., PloS one
Social [51]
306
et al., PloS one
Social [51]
unipartite
Introduction to Mathematical Sociology.
0.00
306
et al., PloS one
Social [51]
unipartite

TABLE I: Details of the ecological and social networks used in the main
document.

| Network Type | Network | Nodes | Dens. | PloS one |
|--------------|---------|-------|-------|----------|
| Ecological   | M49     | 1028  | 0.00  |          |
| Social       | M50     | 306   | 0.01  |          |

[1] B. D. Patterson and W. Atmar, Biological Journal of the Linnean Society 28, 65 (1986).
[2] J. Bascompte, P. Jordano, C. J. Melián, and J. M. Olesen, Proceedings of the National Academy of Sciences 100, 9383 (2003).
[3] In the case of bipartite networks, the comparison must be done between nodes of the same class.
[4] S. Saavedra, F. Reed-Tschesch, and B. Uzzi, Nature 457, 463 (2009).
[5] U. Bastolla, M. A. Fortuna, A. Pascual-García, A. Ferrera, B. Luque, and J. Bascompte, Nature 458, 1018 (2009).
[6] B. Uzzi, American Sociological Review 61, 674 (1996).
[7] S. Bustos, C. Gomez, R. Hausmann, and C. A. Hidalgo, PloS one 7, e49393 (2012).
[8] S. Saracco, R. Di Clemente, A. Gabrielli, and T. Squartini, Scientific Reports 5, 10595 (2015).
[9] J. Borge-Holthoefer, R. A. Baños, C. Gracia-Lázaro, and Y. Moreno, Scientific Reports 7, 41673 (2017).
[10] J. M. Kamilar and Q. D. Atkinson, Proceedings of the National Academy of Sciences 111, 111 (2014).
[11] G. Cimini, A. Gabrielli, and F. S. Labini, PLoS ONE 9, 1 (2014).
[12] M. D. König and C. J. Tessone, Physical Review E 84, 056108 (2011).
[13] M. Bardoscia, G. Luca, G. Livan, M. Marsili, and C. J. Tessone, Journal of Statistical Physics 151, 440 (2013).
[14] M. D. König, C. J. Tessone, and Y. Zenou, Theoretical Economics 9, 695 (2014).
[15] J. Coleman, Introduction to Mathematical Sociology. (Collier-Macmillan: London, UK, 1964).
[16] S. Fortunato, Physics Reports 486, 75 (2010).
[17] R. Guimerà and L. A. N. Amaral, Nature 433, 895 (2005).
[18] J. Borge-Holthoefer and A. Arenas, Entropy 12, 1264 (2010).
[19] S. Fortunato and D. Hric, Physics Reports 659, 1 (2016).
[20] E. Thébault and C. Fontaine, Science 329, 853 (2010).
[21] P. I. Prado and T. M. Lewisohn, Journal of Animal Ecology 73, 1168 (2004).
[22] T. M. Lewisohn, P. Inácio Prado, P. Jordano, J. Bascompte, and J. M. Olesen, PLoS one 13, 174 (2006).
[23] M. A. Fortuna, D. B. Stouffer, J. M. Olesen, P. Jordano, D. Mouillot, B. R. Krasnov, R. Poulin, and J. Bascompte, Journal of Animal Ecology 79, 811 (2010).
[24] C. O. Flores, S. Valverde, and J. S. Weitz, The ISME journal 7, 520 (2013).
[25] J. M. Olesen, J. Bascompte, Y. L. Dupont, and P. Jordano, Proceedings of the National Academy of Sciences 104, 19891 (2007).
[26] M. E. J. Newman, Phys. Rev. E 69 (2004).
[27] M. Almeida-Neto, P. Guimarães, P. R. Guimarães, R. D. Loyola, and W. Ulrich, Oikos 117, 1227 (2008).
[28] C. O. Flores, J. R. Meyer, S. Valverde, L. Farr, and J. S. Weitz, Proceedings of the National Academy of Sciences 108, E288 (2011).
[29] D. Karaboga, An idea based on honey bee swarm for numerical optimization. Techn. Rep. (Technical report-tr06, Erciyes university, engineering faculty, computer engineering department, 2005).
[30] A. Grimm and C. J. Tessone, Applied Network Science 2, 37 (2017).
[31] S. Sobolevsky, R. Campari, A. Belyi, and C. Ratti, Physical Review E 90, 012811 (2014).
[32] M. Meilă, Learning Theory and Kernel Machines , 173 (2003).
[33] R. Guimerà, M. Sales-Pardo, and L. A. N. Amaral, Physical Review E 70, 025101 (2004).
[34] “Web of life: ecological networks database,” http://www.web-of-life.es/.
[35] M. J. Barber, Physical Review E 76, 066102 (2007).
[36] S. Kojaku and N. Masuda, Physical Review E 96, 052313 (2017).
[37] T. P. Peixoto, Physical Review X 4, 011047 (2014).
[38] S. Jonhson, V. Domínguez-García, and M. A. Muñoz, PloS one 8, e74025 (2013).
[39] V. Domínguez-García and M. A. Muñoz, Scientific Reports 5, 8182 (2015).
[40] S. J. Beckett and H. T. Williams, Interface focus 3, 20130033 (2013).
[41] C. Y. J. Leung and J. S. Weitz, Physical Review E 93, 032303 (2016).
[42] A. Davis, B. B. Gardner, and M. R. Gardner, Deep South: A social anthropological study of caste and class (Univ of South Carolina Press, 2009).
[43] M. Sarwat, J. J. Levandoski, A. Eldawy, and M. F. Mokbel, IEEE Transactions on Knowledge and Data Engineering 26, 1384 (2014).
[44] J. J. Levandoski, M. Sarwat, A. Eldawy, and M. F. Mokbel, in Data Engineering (ICDE), 2012 IEEE 28th International Conference on (IEEE, 2012) pp. 450–461.
[45] G. Van de Bunt, M. A. Van Duijn, and T. A. Snijders, Computational & Mathematical Organization Theory 5, 167 (1999).
[46] T. A. Snijders, C. E. Steglich, and G. G. van de Bunt, in Social Networks, 1 (2008).
[47] J. Stohlé, N. Vórin, A. Barrat, C. Cattuto, L. Isella, J.-F. Pinton, M. Quaggiotto, W. Van den Broeck, C. Régis, B. Lina, et al., PloS one 6, e23176 (2011).
[48] J. Forrest and A. Barrat, PloS one 9, e107878 (2014).
[49] R. Guimerà, L. Danon, A. Díaz-Guilera, F. Giralt, and A. Arenas, Physical review E 68, 065103 (2003).
[50] W. W. Zachary, Journal of anthropological research 33, 452 (1977).
[51] B. Klimt and Y. Yang, Proc. Eur. Conf. Machine learning (ECML), 217 (2004).