



Nestedness in complex networks: Observation, emergence, and implications



Manuel Sebastian Mariani ^{a,b,*}, Zhuo-Ming Ren ^c, Jordi Bascompte ^d, Claudio Juan Tessone ^b

^a Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu 610054, PR China

^b URPP Social Networks, University of Zurich, CH-8050 Zurich, Switzerland

^c Alibaba Research Center for Complexity Sciences, Alibaba Business School, Hangzhou Normal University, Hangzhou 311121, China

^d Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

ARTICLE INFO

Article history:

Received 7 November 2018

Received in revised form 16 April 2019

Accepted 18 April 2019

Available online 27 May 2019

Editor: I. Procaccia

Keywords:

Complex systems

Socio-economic networks

Ecological networks

Nestedness

Economic complexity

Emergence

ABSTRACT

The observed architecture of ecological and socio-economic networks differs significantly from that of random networks. From a network science standpoint, non-random structural patterns observed in real networks call for an explanation of their emergence and an understanding of their potential systemic consequences. This article focuses on one of these patterns: nestedness. Given a network of interacting nodes, nestedness can be described as the tendency for nodes to interact with subsets of the interaction partners of better-connected nodes. Known since more than 80 years in biogeography, nestedness has been found in systems as diverse as ecological mutualistic systems, world trade, inter-organizational relations, among many others. This review article focuses on three main pillars: the existing methodologies to observe nestedness in networks; the main theoretical mechanisms conceived to explain the emergence of nestedness in ecological and socio-economic networks; the implications of a nested topology of interactions for the stability and feasibility of a given interacting system. We survey results from variegated disciplines, including statistical physics, graph theory, ecology, and theoretical economics. Nestedness was found to emerge both in bipartite networks and, more recently, in unipartite ones; this review is the first comprehensive attempt to unify both streams of studies, usually disconnected from each other. We believe that the truly interdisciplinary endeavor – while rooted in a complex systems perspective – may inspire new models and algorithms whose realm of application will undoubtedly transcend disciplinary boundaries.

© 2019 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Contents

1. Introduction.....	3
1.1. What are nested networks?	4
1.2. Historical background	6
1.3. How to read this article.....	7
2. Classes of (potentially) nested networks.....	7

* Corresponding author.

E-mail addresses: manuel.mariani@business.uzh.ch (M.S. Mariani), claudio.tessone@uzh.ch (C.J. Tessone).

2.1.	Nested networks in graph theory.....	8
2.1.1.	Nested split graphs	8
2.1.2.	Threshold graphs.....	8
2.2.	Ecological networks.....	9
2.2.1.	Mutualistic networks	9
2.2.2.	Antagonistic networks	10
2.2.3.	Spatial networks.....	10
2.3.	Socio-economic networks	11
2.3.1.	Country-level trade networks	11
2.3.2.	Contractor networks	12
2.3.3.	Interbank networks.....	13
2.3.4.	Spatial networks.....	13
2.3.5.	Communication networks in organizations	14
2.3.6.	Online communication networks and software development.....	14
2.4.	Other classes of networks.....	14
3.	Observing nestedness: Metrics and null models.....	15
3.1.	Metrics to quantify nestedness	15
3.1.1.	Gap-counting metrics	16
3.1.2.	Distance-based metrics.....	17
3.1.3.	Overlap metrics	19
3.1.4.	Eigenvalue-based metrics.....	21
3.2.	Including a null model.....	22
3.2.1.	Nine basic classes of randomization procedures.....	22
3.2.2.	Computing the expected nestedness: Maximum-entropy approach	25
3.2.3.	Statistical tests.....	26
3.3.	Nestedness maximization: Packing the adjacency matrix	27
3.3.1.	Genetic algorithms	27
3.3.2.	Non-linear iterative algorithms	28
3.3.3.	Nestedness temperature minimization: BINMATNEST or fitness-complexity?	31
3.4.	Bottom-line: How to measure nestedness?	31
4.	Relation between nestedness and other systemic properties.....	32
4.1.	Nestedness and other network properties	32
4.1.1.	Nestedness and degree distribution.....	33
4.1.2.	Nestedness and disassortativity	34
4.1.3.	Nestedness and core-periphery structure.....	35
4.1.4.	Nestedness and modularity	37
4.2.	Nestedness and ecological properties	40
4.2.1.	Nestedness and macroecological properties	40
4.2.2.	Spatial networks: Nestedness and beta diversity	41
4.2.3.	Interaction networks: Nestedness and relative species abundance distribution	42
4.3.	Nestedness and economic properties	43
4.3.1.	Individual nestedness contributions and the survival of firms	43
4.3.2.	Nestedness and prediction of economic links	44
4.3.3.	Nestedness maximization and the future development of countries	45
4.3.4.	Nestedness and hidden economic capabilities	46
5.	Emergence of nestedness: Rewiring and formation mechanisms	47
5.1.	Graph-theoretic mechanisms	48
5.1.1.	Threshold model.....	48
5.1.2.	Growing perfectly nested networks with arbitrary degree distribution	48
5.2.	Ecological mechanisms.....	49
5.2.1.	Self-organizing network model	50
5.2.2.	Bipartite cooperation: Specialization and interaction	50
5.2.3.	Fitness maximization based on population dynamics	51
5.2.4.	Nestedness as an evolutionary spandrel: Speciation and divergence	53
5.2.5.	Invasion dynamics with different assemblage rules	54
5.3.	Social and economic mechanisms	54
5.3.1.	Network formation based on centrality maximization	54
5.3.2.	Network formation in a two-stage game	55
5.3.3.	Network rewiring based on social climbing	56
5.3.4.	Emergence of nestedness in the country-product network: Innovation and novelty	57
5.4.	Bottom-line: Why does nestedness emerge in empirical networks?	58
6.	Implications of nestedness for systemic stability and feasibility	58
6.1.	Robustness against structural perturbations	59
6.1.1.	Topological robustness of nested networks	59
6.1.2.	Nestedness maximization and the structural importance of individual nodes	60
6.2.	Stability and feasibility of systems with competition and mutualism in the population dynamics approach	62

6.2.1.	Stability and feasibility of equilibrium points	63
6.2.2.	Stability and feasibility in a two-species competition model	64
6.2.3.	Stability and feasibility in models with competition and mutualism	65
6.2.4.	The impact of nestedness on feasibility and global stability.....	68
6.3.	Local stability of mutualistic systems in the random-matrix theory approach.....	70
6.3.1.	Stability-complexity relation: May's result for random matrices.....	70
6.3.2.	Stability of mutualistic networks	71
6.3.3.	Stability of nested mutualistic networks	71
6.4.	Bottom-line: What is the impact of nestedness on robustness, feasibility, and stability?	72
7.	Observing nestedness at the mesoscopic scale	73
7.1.	Quantifying the degree of nestedness of a subgraph	74
7.2.	Finding the largest nested subgraph	75
7.3.	Detecting in-block nestedness	75
7.3.1.	Nestedness of network modules	75
7.3.2.	Detecting in-block nested structures	76
7.3.3.	Detecting multiple core-periphery pairs	77
8.	Outlook	80
	Acknowledgments	80
	Appendix A. Software and relevant datasets	81
A.1.	Software for nestedness and network analysis	81
A.1.1.	Metrics and null models	81
A.1.2.	Nested network generation	81
A.1.3.	Community detection and nestedness at the mesoscopic scale	81
A.2.	Datasets of socio-economic and ecological networks	81
A.2.1.	World trade data	81
A.2.2.	Ecological spatial networks	81
A.2.3.	Ecological interaction networks	81
	Appendix B. Network-based metrics of node importance in bipartite networks	82
	References	83

1. Introduction

Network science constitutes one of the pillars of the modern science of complexity. Broadly speaking, network science can be defined as the “study of the collection, management, analysis, interpretation, and presentation of relational data” [1]. Given a complex system with many interacting components, a suitable network representation simplifies its analysis by describing it – in the simplest portrayal – as a collection of nodes and links connecting them. The working hypothesis is that while the obtained network represents a simplification of the complex system it was derived from, it carries enough information to allow us to understand the functioning of the system and the emergence of collective phenomena from the interactions of its constituents. Leveraging its inherently interdisciplinary nature, network science borrows ideas and tools from a variety of research fields, especially those where networks (also, especially in mathematics and computer science, termed *graphs*) have been long studied: graph theory [2], social science [3], economics [4], scientometrics [5,6], computer science [7], and ecology [8,9], among others.

Perhaps one of the most intriguing features of real networks is the existence of common structural and dynamical patterns that are found in a large number of systems from various domains of science, nature, and technology. The pervasiveness of structural patterns in systems from various fields makes it possible to analyze them with a common set of tools [10]. A popular example of such widespread structural patterns is the heavy-tailed degree distribution [11] (i.e., the distribution of the number of connections per node). Power-law degree distributions (often termed as “scale-free”) have been reported in a wide variety of different systems, ranging from social and information networks to protein-protein interaction networks (see [12] for a dedicated book). The ubiquity of heavy-tailed degree distributions has motivated a large number of studies aimed at unveiling plausible mechanisms that explain their emergence [10,13], understanding their implications for spreading [14,15], network robustness [16], synchronization phenomena [17,18], etc.

This review article focuses on one of these widely-observed (and, as a result, widely studied and debated) network structural patterns: *nestedness*. Informally speaking, nestedness refers to a hierarchical organization where the set of neighbors of a node is a subset (superset) of the neighbors of lower (larger) degree. Originally conceived [19,20] and discovered [21] in biogeography where it was found in the spatial distribution patterns of species [21,22], nestedness has been found in a wide variety of systems, including ecological interaction networks [23], trade networks [24–26], inter-organizational networks [27,28], firm spatial networks [29,30], interbank payment networks [25,31], social-media information networks [32], among others. Nestedness has been observed in two distinct families of networks. On the one hand, *bipartite networks* (in which nodes of two categories are connected in some way), like mutualistic networks in ecology. On the other hand, more recently, nestedness has been found also in unipartite networks, i.e., those where all the nodes are of the same nature, like interacting agents in economics. The two streams of research have run mostly in

parallel, paying little or no attention to each other. To the best of our knowledge, for the first time, this review provides an in-depth discussion of both research areas.

The widespread occurrence of a given pattern in systems of different nature naturally leads to a number of questions related to the fundamental mechanisms behind its emergence and its implication for the systems' functioning and preservation. Based on the massive amount of works on nestedness, this review will address several questions: Which classes of networks exhibit a nested architecture? What are the possible mechanisms behind the emergence of nestedness in ecological and socio-economic networks? Is the emergence of nestedness the result of an optimization process where constituents tend to maximize their fitness [33], or is it a consequence of the assemblage rules of the system [34,35]? Which mechanisms are responsible for the emergence of nestedness in social systems and trade networks? [25,36] Is nestedness beneficial or harmful for the stability [37] and the persistence [38] of a given ecological community? How is nestedness related to other widely observed network properties, such as modularity [39,40] and core-periphery structure [41]? Given a network, is nestedness a global property, or a property that is only respected by specific subsets of the network? [40,42–44] How can we exploit nestedness for predicting future properties of economic systems, such as the gross domestic product of countries [45] and the appearance of a new firm at a given geographic location [29]?

Admittedly, some of these questions have been controversial in the literature, and sometimes, contradictory results have been supported by different studies. A complete coverage of all existing viewpoints on each topic would be impossible to attain. However, when necessary, we will try our best to present the most relevant findings that have been used to support competing claims.

This review article aims at providing the reader with a cross-disciplinary perspective on nestedness, the subtleties associated with its observation, the possible mechanisms behind its emergence, and its implications. Driven by these questions, our review is an interdisciplinary effort where we collected results from network science, statistical physics, ecology, economics, social sciences, among others, which makes our review fundamentally different from recent ones [46–49] and books [9] which have focused exclusively on the structure and dynamics of ecological networks. We hope that this effort will allow researchers who have been interested in nestedness to learn the methodologies and findings of scholars from other fields, which might accelerate the progress of research on this area and, at the same time, reveal similar organizing principles across disciplines [27].

A substantial part of the methodologies detailed in this article come from statistical physics and dynamical systems: population dynamics [38,50], network formation models [36,51], dynamical system stability analysis [37], fitness optimization techniques [33,40,44], prediction based on dynamical systems [45]. Other methodologies and results come from theoretical ecology [34,35,38], theoretical economics [25,52], and graph theory [53,54]. We will always try to provide sufficient background information for all the methodologies and problems presented throughout the review.

1.1. What are nested networks?

Before detailing the historical developments of nestedness, we provide the definition of perfectly nested networks together with the basic ideas behind nestedness analysis. In abstract terms, we refer here to systems that are composed of many constituents. Said elements are not isolated but they interact with (or are related to) others. If these interactions can be described as dyadic, the system admits a representation in terms of a *network* [4,10,11]: its constituents are represented as *nodes* (or *vertexes*) and their mutual interactions (or relations) as *links* (or *edges*, or *bonds*).¹ The set of nodes connected to a certain node constitute its *neighborhood*, and the number of such connections is referred to as its *degree*.

For example, if the network under consideration is a friendship network, two individuals are linked when they have amity between them. In a trade network, whose constituents are countries, an edge is present if there exists a trade relationship between two given countries. In the examples above the elements of the system (individuals or countries), while heterogeneous, are of the same nature; it is then said that the network is *unipartite*. There are other networks whose constituents can be divided into two distinct classes. For example, in a habitat, a set of pollinator species and a set of flowering plants are mutually related if the former helps in the pollination (and feeds from) the latter. In this case, relations connect agents of a distinct nature. This construct is called a *bipartite* network.

A *perfectly nested network* is formally defined in the following terms: consider any two nodes i and j ; if the degree of i is smaller than the degree of j , then the neighborhood of i is contained in the neighborhood of j . This definition is generally valid for unipartite and bipartite networks with a single proviso: for the second group, the nodes to be compared must belong to the same class. Any network here considered accepts a representation in terms of a binary adjacency matrix A whose elements A_{ij} are equal to 1 (0) if the two vertices are connected (not connected). For the particular case of nested networks, the adjacency matrix has a very peculiar shape (see lower row in Fig. 1): after re-ordering its rows and columns by degree, there exists a monotonic separatrix above which all the elements are equal to one, and zero below.²

¹ This is not the only way to represent relational data. For example, temporal [55,56] and multilayer [57,58] networks provide with more sophisticated representations that allow us to take into account temporal effects and multiple types of interactions, respectively.

² Strictly speaking, in bipartite network analysis, the matrix that connects nodes of different type is typically referred to as biadjacency matrix [59] or incidence matrix [11], whereas the adjacency matrix still refers to the matrix whose elements represent all possible pairs of nodes – the elements that connect nodes of the same type are, by definition, all equal to zero. To simplify the discussion, when we will present methods for bipartite networks that apply equally well to unipartite networks, we will refer to the biadjacency matrix as the adjacency matrix. This abuse of terminology has no harmful consequences as the elements of the adjacency matrix that connect nodes of the same type are all equal to zero.

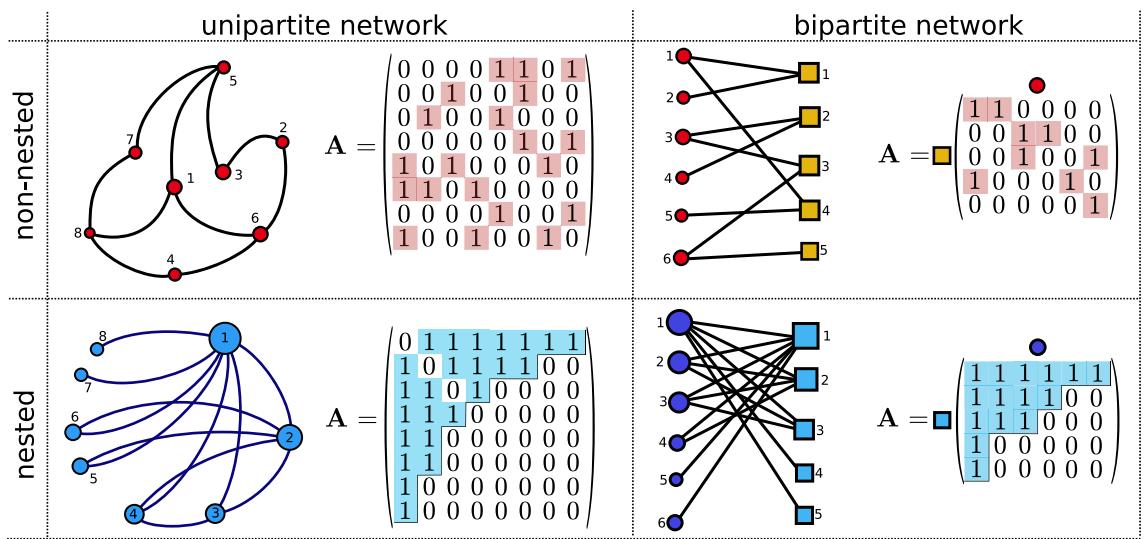


Fig. 1. Simple examples of nested and non-nested networks in both cases: unipartite and bipartite networks. The adjacency matrices of the perfectly nested networks (bottom panels) exhibit a clear “triangular” shape: a separatrix partitions the matrices into unambiguously-defined filled regions (i.e., regions that only contain ones) and empty regions (i.e., regions that only contain zeros). In this review, we refer to these kinds of matrices as perfectly nested matrices.

The first observation is that this definition can accommodate topologies as different as a star network (a highly centralized topology where one hub is connected to all the other nodes) and a fully-connected network (where each node is connected with everyone else). It is also clear, however, that this is a very stringent network topology which is seldom (if ever) observed in real-world networks. Therefore, *a priori*, this kind of network structure may seem a kind of curiosity. However, research has shown that networks showing this property abound in a wide variety of research fields, and its definition respected to a large extent: the vast majority of links are located in nested neighborhoods as defined above. Asking “how much of a network structure respects the property of nestedness” gave rise to several measures that aim at quantifying the *degree of nestedness* of a given network, and several null models to quantify the statistical significance of the observed levels of nestedness. In this review, when we say that a network is “nested”, we mean that it exhibits a level of nestedness (according to a given metric for nestedness) that cannot be explained by a reasonable null model.

From this definition, it is evident that metrics for nestedness and null models used to produce random networks that respect given constraints are of vital importance [60–62]. Different metrics and null models may be responsible for the validation or rejection of the hypothesis of nestedness significance, and different combinations of metrics and null models can lead to different conclusions [63]. Therefore, a proper definition of null models is crucial, and far from trivial: null models purely based on preserving simple topological properties can be either too restrictive – when the full degree sequences must be exactly preserved in bipartite networks, for example – offering little freedom from the observed network [61,63,64]; or too lax – like requiring only that the network density is preserved – in which only rarely the null hypothesis that the empirical degree of nestedness in a network is significant can be rejected [48,65–67]. These issues will be discussed in Sections 3.2 and 4.1.1.

The notion of nestedness introduced above is based on degree. Unless otherwise stated, when we refer to nestedness throughout this review, we refer to the nestedness by degree defined above. One can also take a different perspective and ask whether a network exhibits a significant level of nestedness when pairwise node orderings build on a node-level scalar property different than the degree – e.g., species abundance [68] and islands’ area [69]. In other words, one investigates whether the network’s adjacency matrix exhibit a significant degree of nestedness once its rows and columns are sorted by a given property s . This consideration leads to two additional notions of nestedness that have been studied:

- *Nestedness by other scalar properties.* Consider a node-level scalar property s . One can re-define nestedness based on property s : the neighborhoods of two given nodes i and j such that $s_i < s_j$ are nested by property s if and only if the neighborhood of i is a subset of the neighborhood of j . The network is highly nested by property s if many of its pairs of nodes’ neighborhoods are nested by s . In other words, a network is highly-nested by s if, after ordering the adjacency matrix’s rows and columns by s , the resulting matrix exhibits a clear separation between its filled and its empty region.
- *Maximal nestedness.* One seeks to find the ordering of rows and columns that maximizes the level of nestedness of the network. Scholars have especially aimed to find maximal-nestedness structures in the context of nestedness temperature analysis of ecological networks [22,70] (see Section 3.1.2).

Based on the definitions provided above, one can compute, for instance, the level of nestedness of an ecological mutualistic system based on species abundance, and compare the resulting nestedness with the maximal nestedness that can be achieved by the system [68]. In biogeography, one can study the nestedness of a species-island network by both degree and island-level environmental variables, including area and distance from the mainland — see [69] for a comparative analysis.

Besides, nestedness-maximizing ranking algorithms have found important applications for the prediction of the future development of countries [24,45] and the identification of the most vulnerable constituents in ecological networks [71] and world trade networks [72]. When defining existing metrics for nestedness (Section 3), we will specify which aspect of nestedness they capture. It is worth noticing that metrics for nestedness that take as input a pairwise ordering of the nodes can be applied to measure both nestedness by degree and nestedness by other scalar properties. On the other hand, algorithms that seek to maximize nestedness (like those used to minimize nestedness temperature [22,70,73]) cannot be used to evaluate the level of nestedness by a specific node-level property; however, one can compare the resulting maximal nestedness with the degree of nestedness by a property of interest [68].

1.2. Historical background

While the concept of nestedness was precisely formulated by Patterson and Atmar [21], the preceding literature offers studies [19,20] where the notion of nestedness was already implicitly discussed. In particular, in his book *Zoogeography* [20], Darlington described the potential effect of spatial distance on the patterns of distribution of species on islands. He considered two possible mechanisms of animal over-water dispersal when moving from the continent to islands. He noticed that other things being equal, individuals of a given species might be expected to populate first the nearest island, and then move to the following nearest islands by crossing water gaps of minimal length. This occupation process, referred to as *immigrant pattern* by Darlington [20] and *selective immigration* by subsequent studies [48,74], naturally results in a nested structure where the species found on the farthest islands (i.e., the species with the higher dispersal rate [74]) are also found in the nearest islands (see Fig. 57 in [20]).

Besides, between the 70 s and the 80 s, ecologists became increasingly interested in the spatial distribution patterns of species and statistical null models to assess their statistical significance³ [65,75–77]. It was not until 1986 that the notion of *nestedness* was introduced together with the first metric to measure it. This was done by Patterson and Atmar [21] who found that several species spatial distribution patterns exhibit a highly nested structure that cannot be explained by null models. In a subsequent study, the same authors introduced the popular nestedness temperature [22], a metric for nestedness which has been applied in a large number of papers in ecology [48].

In parallel, graph theory had started investigating *threshold graphs* and *nested split graphs*, which constitute classes of graphs that are equivalent to perfectly nested graphs (see Section 2.1). Remarkably, the term “nested bipartite graph” together with the first mathematical definition of a perfectly nested bipartite graph already appeared in a paper [78] published by Hering in 1971. Rigorous results for threshold graphs and nested split graphs are reviewed in the 1995 book by Mahadev and Peled [79].

In 2003, Bascompte et al. [23] brought the concept of nestedness to ecological interaction networks by analyzing 25 plant–pollinator networks and 27 plant–frugivore networks, finding that most of them exhibit a degree of nestedness that cannot be explained by degree-preserving null models. Motivated by this finding, scholars have aimed at uncovering the basic mechanisms that are potentially responsible for the emergence of nestedness (see Section 5.2). Besides, a large number of works have attempted to understand how nestedness impacts both the structural and the dynamical stability of the system. Scholars have investigated this relation by means of simulating co-extinction cascades [71], numerical simulations and analytic results on models of mutualistic and competitive interactions [38,50], random-matrix theory [37]. We refer to Section 6 for the details.

Stimulated by works on ecological networks, interest in nestedness analysis has arisen in socio-economic networks as well. In 2007, Soramäki et al. [31] found that the interbank payment network is organized in a topology with a densely connected core and a periphery that is only connected with the core: a structure that is compatible with nestedness. Saavedra et al. [27] found that manufacturer–contractor networks exhibit a nested structure, and proposed a parsimonious model to explain its emergence. König et al. [25] found that four economic networks are significantly nested: the Fedwire network of settlements, Austrian inter-bank network, the world trade network and the worldwide arm trade network. Besides, they proposed a network formation model to explain its emergence.

As for bipartite trade networks, Saracco et al. [26] found that when analyzing the bipartite country–product trade network by means of standard nestedness analysis tools, a decrease of nestedness took place before the 2007–2008 financial crisis. In parallel, Tacchella et al. [24] have designed a non-linear ranking algorithm to quantify the competitiveness of countries and the sophistication of products. In fact, the algorithm sorts rows and columns of the network’s adjacency matrix in such a way that the nested architecture of the system is revealed [24,71], and it often outperforms genetic algorithms from the ecological literature in maximizing nestedness [73].

³ In this respect, it is surprising that a network null model that preserves the individual nodes’ degrees, typically referred to as *configuration model* in network science [11], was already designed in 1979 by Connor and Simberloff [65].

1.3. How to read this article

The main goal of this article is to review three main aspects of nestedness: its observation, the modeling of its emergence, and its implications. We detail below which Sections specifically deal with each of these three aspects:

- **Observing nestedness.** Section 2 provides an overview of the classes of real networks of interest for this review. Section 3 addresses the questions: How to measure the level of nestedness of a given system? How to choose a suitable null model to assess the statistical significance of the observed level of nestedness? How to maximize the degree of nestedness of a given network by properly ranking its nodes? Section 7 surveys the methods that aim to quantify nestedness at a mesoscopic scale, which shifts the question from assessing whether a given system is nested to finding sub-components of the system such that the nodes that belong to the found components exhibit a nested interaction topology.
- **Modeling the emergence of nestedness.** Section 5 deals with possible mechanisms to explain the emergence of nestedness that have been identified in graph theory, ecology, and economics. Proposed mechanisms range from optimization principles that postulate that nodes aim to maximize their fitness [33] or their centrality [25,36,80,81] when choosing their interactors, to stochastic processes where nestedness is the outcome of duplication and link randomization processes [34]. No agreement on the main mechanisms behind the emergence of nestedness has been reached yet.
- **Implications of nestedness.** Section 6 deals with both the implications of nestedness for network robustness against targeted attacks, for the stability and feasibility of equilibrium points of mutualistic dynamical processes on interaction networks. In this context, *stability* refers to the ability of the system to return to its original equilibrium state after a small perturbation (*local* stability) or after a perturbation of any magnitude (*global* stability); *feasibility* refers to the existence of equilibrium points such that all species are represented by at least one individual. The latter problem is critical to the assessment of the impact of nestedness on the co-existence of species in ecological systems. The variety of methods adopted in the literature to tackle these problems has led to variegated conclusions.

The Outlook section will point to the current challenges in research in nestedness as well as the most promising unexplored directions in the topic.

Due to its interdisciplinary nature, our review encompasses methodologies and findings from diverse fields. For readers interested in different aspects of nestedness, we provide the following road-map:

- A reader interested in the **ecological** aspects of nestedness analysis may consider reading the following sections: Section 2.2 for a short overview of the area; 3 to learn about the technical aspects of measuring nestedness; 4.2.3 to learn about the relation between nestedness and other properties; 5.2 and 5.4 for the main mechanisms that have been proposed in the ecological literature to explain the observed nestedness of interaction networks; 6 for the implications of nestedness for systemic robustness, stability, and feasibility; 8 for an overview of current challenges and open directions in research on nestedness.
- A reader interested in the **socio-economic** aspects of nestedness analysis may consider reading the following sections: Section 2.3 for a brief overview of socio-economic networks of interest; 3 to learn about the technical aspects of measuring nestedness; 4.3 to learn about the applications of nestedness to predictive problems in socio-economic systems; 5.3 to learn about possible social and economic mechanisms that can explain the emergence of nestedness in socio-economic systems; 8 for an overview of current challenges and open directions in research on nestedness.
- A reader interested in the **methodological** aspects of nestedness analysis may consider reading the following Sections: 2.1 to learn about the equivalence between perfectly nested networks, nested split graphs, and threshold graphs; 3 to learn about the technical aspects of measuring nestedness; 5.1 to learn how to generate perfectly nested networks with different degree distributions; 7 to learn the technical aspects and main implications of detecting nestedness at the mesoscopic scale.
- A reader who is already familiar with the literature of nestedness in ecological systems and would like to catch up with the **most recent developments** in the topic may consider reading the following Sections: Section 5 to discover recent mechanisms that have been proposed to explain the emergence of nestedness in interaction networks; 6 to discover the recent developments on the implications of nestedness for systemic robustness, feasibility, and stability; 7 to learn the technical aspects and main implications of detecting nestedness at the mesoscopic scale; 8 for an overview of current challenges and open directions in research on nestedness.

The notation of the review for unipartite and bipartite networks is shown in Table 1 and Table 2, respectively.

2. Classes of (potentially) nested networks

In this Section, we present classes of graph-theoretical, ecological, economic and social systems where a nested architecture has been found. For each of these research fields, we provide a classification of the types of networks involved, and we mention some of the most representative papers that pointed out their nested architecture. The main goal of this Section is to provide readers with a brief introduction to the specific language of the systems studied in this review, without aiming to be exhaustive. For each of the systems presented below (see Table 3 for a summary), wherever possible, we point to Web repositories where datasets can be found (see Appendix A).

Table 1

Notation for binary unipartite networks. Mathematically equivalent concepts are expressed through different terms in different disciplines. In this review, we will mostly use the network science nomenclature, yet we will sometimes switch to the ecology language when presenting results in ecology.

Symbol	Network science	Graph theory	Ecology
\mathcal{G}	Network	Graph	Community
$i \in \{1, \dots, N\}$	Node	Vertex	Species
N	Network size	Graph order	Community size
(i, j)	Link, edge, bond	Edge, arc	Interaction
$A = \{A_{ij}; A_{ij} = 1 \text{ iff } (i, j) \text{ is observed}\}$	Adjacency matrix	Adjacency matrix	Interaction matrix
$E = \sum_{ij} A_{ij}$	Number of links/edges	Graph size	Number of pairwise interactions
$\rho = 2E/(N(N - 1))$	Connectance or Density	Graph density	Fill
$\mathcal{N}_i = \{j A_{ij} = 1\}$	Neighborhood	Neighborhood	Set of interactors
$k_i = \sum_j A_{ij}$	Degree	Degree, valency	Marginal total
$\langle k \rangle = N^{-1} \sum_j k_j = 2E/N$	Average degree	Average degree	Average number of interactions per species
Node i with $k_i \gg \langle k \rangle$	High-degree node, hub	High-degree node	Generalist
Node i with $k_i \sim 1$	Low-degree node	Low-degree node	Specialist
$O_{ij} = \sum_k A_{ik} A_{jk}$	Common neighbors	Common neighbors	Overlap
\mathcal{E}_i	i 's Community	i 's Cluster	i 's Compartment

2.1. Nested networks in graph theory

This article will focus on the “physical” aspects of nestedness, including the mechanisms that lead to its emergence and its implications for dynamical processes on networks. While rigorous results from graph theory will not constitute a central topic, it is instructive to acknowledge the deep connection between nested networks and classes of graphs that have been widely studied: threshold and split graphs [79]. The fact that many rigorous results have been obtained for these two classes of graphs [79] might inspire their application to the analysis of nested ecological and economic networks. Rigorous theorems on graph theory [82] have been used, for example, by Staniczenko et al. [83] to motivate the spectral radius of the adjacency matrix as a possible metric for nestedness (see Section 3.1.4). For the sake of completeness, we also mention that triangulated planar graphs are also nested networks [84] and can be leveraged for hierarchical clustering of empirical data [85]. In the following, we narrow down our focus to unweighted unipartite networks (graphs), and define nested split and threshold graphs.

2.1.1. Nested split graphs

Split graphs have been originally introduced and studied by Földes and Hammer [86] and, independently, by Chernyak and Chernyak [87]. A graph \mathcal{G} can be referred to as a split graph if and only if its nodes can be partitioned into a *clique* (or *dominating set*) \mathcal{K} and a *stable set* (or *independent set*) \mathcal{S} – in this case, one denotes the split graph as $\mathcal{G}(\mathcal{K}, \mathcal{S})$. In the graph-theoretical language, a *clique* [88] is a subset of nodes such that every pair of nodes that belong to the clique is connected through a link; a *stable set* [89] is a subset of nodes such that no two nodes are connected through a link.

A split graph $\mathcal{G}(\mathcal{K}, \mathcal{S})$ is referred to as *nested split graph* [90] if and only if the neighborhoods of the nodes that belong to the stable set \mathcal{S} are nested: given two nodes i and j that belong to \mathcal{S} , if $k_i < k_j$, then the neighborhood of i is included in the neighborhood of j .⁴ By definition, a nested split graph is also a perfectly nested network. Such equivalence will be exploited in Section 5.1 to introduce a generative algorithm for perfectly nested networks. The fingerprint of a nested split graph is a *stepwise adjacency matrix* [91]: a symmetric binary matrix A whose elements A_{ij} satisfy the property that if $i < j$ and $A_{ij} = 1$, then $A_{hk} = 1$ if $h < k < j$ and $h \leq i$. A stepwise matrix exhibits a peculiar “triangular” shape (see Fig. 1).

2.1.2. Threshold graphs

Importantly, nested split graphs are equivalent to an important class of graphs: the threshold graphs [79]. *Threshold graphs* were introduced in the 70 s by Chvátal and Hammer [92]. A graph is referred to as a threshold graph if and only if there exist non-negative real numbers $\{w_i\}, \theta$ such that a subset \mathcal{U} of the nodes is stable if and only if its total weight $w(\mathcal{U}) := \sum_{i \in \mathcal{U}} w_i$ does not exceed the threshold θ : $w(\mathcal{U}) \leq \theta$. One can prove (Theorem 1.2.4 in the book [79]) that given a graph \mathcal{G} , the property that \mathcal{G} is a threshold graph is equivalent to the property that \mathcal{G} is a nested split graph. Therefore, nested split graphs, threshold graphs, and perfectly nested networks are equivalent.

⁴ In graph theory, one says that vertex j dominates vertex i [79].

Table 2

Notation for binary bipartite networks. Mathematically equivalent concepts are expressed through different terms in different disciplines. In this review, we will mostly use the network science nomenclature, yet we will sometimes switch to the ecology language when presenting results in ecology.

Symbol	Network science	Graph theory	Ecology	Biogeography
\mathcal{G}	Network	Graph	Community	Region, archipelago
\mathcal{R}, \mathcal{C}	Two disjoint sets of nodes	Two disjoint sets of vertices	Set of active and passive species	Set of patches and species
$i \in \{1, \dots, N\}$	Class- \mathcal{R} 's node, Row-nodes	Class- \mathcal{R} 's vertex	Active species	Species
$\alpha \in \{1, \dots, M\}$	Class- \mathcal{C} 's node, Column-node	Class- \mathcal{C} 's vertex	Passive species	Patches, islands, sites
$S = N + M$	Network size	Graph order	Community size	
(i, α)	Link, edge, bond	Edge, arc	Interaction	Presence
$A = \{A_{i\alpha}; A_{i\alpha} = 1 \text{ iff } (i, \alpha) \text{ is observed}\}$	(Bi-)Adjacency or Incidence matrix	(Bi-)Adjacency or Incidence matrix	Interaction matrix	Presence-absence matrix
$E = \sum_{i,\alpha} A_{i\alpha}$	Number of links/edges	Graph size	Number of pairwise interactions	
$\rho = E/(NM)$	Connectance, Density	Graph density	Fill	Fill
$\mathcal{N}_i = \{\alpha A_{i\alpha} = 1\}$	i 's Neighborhood	i 's Neighborhood	Set of i 's partners	Set of sites where i is present
$\mathcal{N}_\alpha = \{i A_{i\alpha} = 1\}$	α 's Neighborhood	α 's Neighborhood	Set of α 's partners	Set of α 's present species
$k_i = \sum_\alpha A_{i\alpha}$	Degree	Degree, valency	Row's marginal total or Active species' degree	Species frequency
$k_\alpha = \sum_i A_{i\alpha}$	Degree	Degree	Column's marginal total or Passive species' degree	Site richness
$\langle k^R \rangle = N^{-1} \sum_j k_j$	Average degree of class- \mathcal{R} nodes	Average degree of class- \mathcal{R} nodes	Average rows' marginal total	Average species frequency
$\langle k^C \rangle = M^{-1} \sum_\alpha k_\alpha$	Average degree of class- \mathcal{C} nodes	Average degree of class- \mathcal{C} nodes	Average columns' marginal total	Average site richness
Node i with $k_i \gg \langle k^R \rangle$	Hub	Hub	Generalist	Ubiquitous species
Node i with $k_i \sim 1$	Low-degree node	Low-degree node	Specialist	Rare species
$O_{ij} = \sum_\alpha A_{i\alpha} A_{j\alpha}$	Common neighbors	Common neighbors	Rows' Overlap	Rows' Overlap
$O_{\alpha\beta} = \sum_i A_{i\alpha} A_{i\beta}$	Common neighbors	Common neighbors	Columns' Overlap	Columns' Overlap
\mathcal{E}_i	i 's Community	i 's Cluster	i 's Compartment	

2.2. Ecological networks

As ecologists are typically interested in studying how species interact with one another and with their surroundings, networks naturally emerge as a powerful tool to study ecological systems. We introduce here the basic concepts of three broad classes of ecological networks [9]: mutualistic networks, antagonistic networks, and spatial networks.

2.2.1. Mutualistic networks

Mutualistic systems are typically represented as bipartite networks composed of two kinds of nodes, corresponding to animal and plant species. In such networks, *mutualism* is the key element: the nodes of one class benefit from the interactions with the nodes of the other class. Different types of mutualism exist in nature. In line with [9], in this review, we focus on two main types of mutualistic interactions: pollination and seed-dispersal. Both mechanisms involve the dissemination of an agent (pollen or seeds) performed by animals, in exchange for nutrients provided by the plant [93]. Yet, there are fundamental differences between the two types of interactions: for example, pollen has to be transferred to a very specific location (the stigma of a conspecific flower), whereas a similar spatial target is not necessary for seeds. We refer to [93] for an insightful discussion of analogies and differences between the two mechanisms. For our purposes, we consider mutualistic networks as bipartite networks where animals (“active species”) interact with plants (“passive species”).

Table 3

Main networks of interest in this review, together with the interpretation of their nodes and links. Networks where the row- and column-nodes are the same are unipartite.

Research field	System type	Row nodes	Column nodes	Links
Ecology	Spatial network, region	Species	Island/site	Species occupy islands/sites
Ecology	Pollination network	Pollinator animal	Plant	Pollinator pollinates plant
Ecology	Seed-dispersal network	Frugivore animal	Plant	Frugivore disperses plant's seed
Economics	Interbank payment	Bank	Bank	Bank pays bank
Economics	Manufacturer-contractor	Manufacturer	Contractor	Manufacturer outsources activity to contractor
Economics	Trade	Country	Country	Country trades with country
Economics	International trade	Country	Product	Country exports product

The study of mutualistic systems has a long and fascinating history in ecology.⁵ Nestedness in both pollination and seed-dispersal mutualistic networks has been found in 2003 by Bascompte et al. [23]. That finding has spurred a large number of works that have aimed at understanding both the possible mechanisms behind the emergence of nestedness (see Section 5.2), and the potential implications of nestedness for a given mutualistic system (see Section 6). Some of the conclusions on the impact of nestedness on systemic stability are based on population dynamics models that describe how the abundances of plant and animal species are affected by the abundances of their interactors, given a topology of interactions. These models of mutualism are described in Section 6.2.3.

2.2.2. Antagonistic networks

In contrast with mutualistic interactions, antagonistic pairwise interactions benefit one species to the detriment of the other. We focus here on two classes of antagonistic networks: food webs and host-parasite networks.

Food webs. While early studies of food webs date back to the first half of the XX century [9], analysis of the structure and stability of food webs intensified in the 70 s, also fueled by May's fundamental theoretical work on the stability-complexity relation [94] (see Section 6.3.1). The simplest type of food web is a unipartite directed network where a directed link from species i to species j means that species j preys on species i [11]. If a set of species of interest preys on and is predated by the same sets of species, mostly, one can coarse-grain the network by representing the set of species of interest as a single node, referred to as trophic species [11]. This unipartite directed network is typically referred to as *community food web*. Besides, ecologists have considered subsets of a complete food web by selecting, for example, a specific type of "resource" species (e.g., plants), and all the species (e.g., herbivores) that consume the resource species. The resulting bipartite networks are typically referred to as resource-consumer [23] or source/sink networks [11,47].

The ecological literature has reported mixed findings on the nestedness of resource-consumer networks. Bascompte et al. [23] found that their level of nestedness is significantly lower than that of mutualistic networks. This finding has not been confirmed by a subsequent study by Kondoh et al. [95], who found that the degree of nestedness of trophic networks is comparable to that of mutualistic networks. As the nestedness of mutualistic networks has been attracted substantially more attention than that of food webs, we mainly focus on mutualistic networks in this article.

Host-parasitoid networks. Parasitoids represent a special type of predators: they lay their eggs inside, on the surface of or near their hosts, and the feeding larvae use their host as food [9]. This kind of interaction has been widely studied in ecology. As being based on antagonistic interactions, one would expect host-parasitoid networks to exhibit substantially different structural patterns compared to mutualistic networks [9,96]. Yet, some studies [97,98] found that host-parasite networks can exhibit a nested structure, and scholars have investigated possible dynamical mechanisms that could lead to its emergence [99].

2.2.3. Spatial networks

From a historical standpoint, following the seminal work by Patterson and Atmar [21], spatial networks constitute the first class of systems where nestedness has been extensively studied. In particular, scholars have been interested in measuring the degree of nestedness of bipartite species distribution networks: such networks are composed of two classes of nodes, one representing species, and the other one representing habitat patches (such as islands). In this review, spatial networks will be often discussed in the Sections concerning distance-based metrics for nestedness (Section 3.1.2) and null models (Section 3.2), yet surveying all the mechanisms that have been proposed to explain the nested structure of spatial networks falls outside the scope of this article.

⁵ We refer to Chapter One of the book [9] for a historical overview of studies of mutualism in ecology.

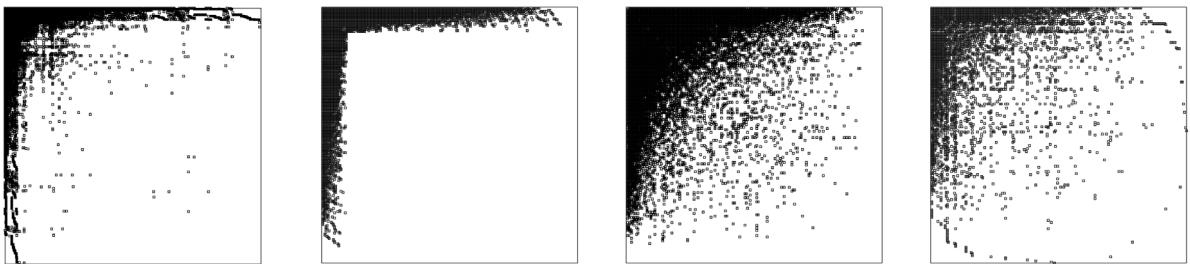


Fig. 2. The adjacency matrices of four economic unipartite networks; their rows and columns have been sorted by degree. From left to right, the Austrian banking network; the global network of banks obtained from the Bank of International Settlements (BIS) locational statistics, the world trade network and the arm trade network. All these networks exhibit a significant level of nestedness.

Source: Reprinted from [25].

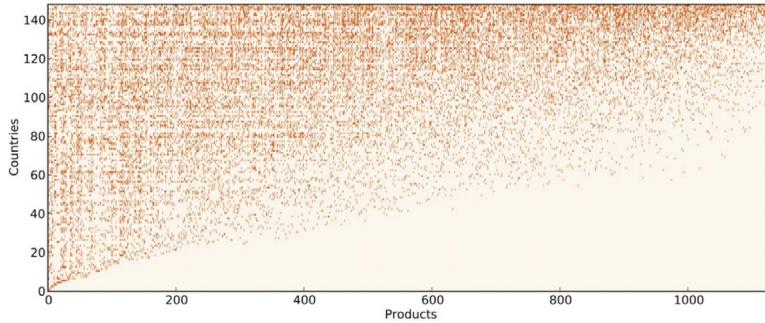


Fig. 3. The (bi-)adjacency matrix of the bipartite country–product network (BACI dataset, year 2010), where the rows (countries) and columns (products) have been sorted according to the fitness–complexity algorithm (see Section 3.3.2). The existence of highly-diversified countries (top rows) contradicts the standard view that the wealthiest countries should focus on trading few products with a high degree of specialization [100]; instead, the data show that the most developed countries tend to diversify their export baskets. This finding is the basis for capability-based models that aim to explain the topology of the country–product network (see Section 4.3.4).

Source: Reprinted from [24].

2.3. Socio-economic networks

Nestedness has been long-studied in ecology, at the point that it is widely regarded as a central structural trait for ecological network analysis. At the same time, it has recently attracted interest in the analysis of social and economic systems [24,25,29,32].

Both unipartite and bipartite socio-economic networks can exhibit significant levels of nestedness. Significant levels of nestedness have been found in various classes of socio-economic unipartite networks (see Fig. 2 for an illustration), including: interbank networks [25], the country–country world trade network [25] and the worldwide arm trade network [25]. Bipartite networks that exhibit significant nestedness include the country–product export network [24,26,29] (see Fig. 3 for an illustration), manufacturer–contractor networks [27], firm–location spatial networks [29], user–meme networks in social media platforms [32].

This is particularly intriguing because it suggests that socio-economic and ecological systems can exhibit qualitatively comparable structures, which can be useful both to understand their common robustness properties, and to design effective strategies to prevent catastrophic events [101].

2.3.1. Country-level trade networks

World trade datasets usually feature the volumes of exports of products between countries over multiple years. From the raw data, one can then build a country–country trade network or a country–product bipartite network, as detailed below. We refer to Appendix A for the links to publicly available world trade datasets.

Unipartite trade networks. The structure of the International Trade Network has been widely studied, both from the country–country and from the country–product perspective. Given publicly available data on the export of products from one country to another, it is possible to construct unipartite country–country networks where nodes represent countries, and links between countries represent the existence of a trade exchange of a product between the two (binary network representation) [102], or the total volume of export between the two countries (weighted network representation) [103]. Scholars have investigated several structural properties of unipartite trade networks, including their community structure [104], their centralization [105], their clustering coefficient [102].

Importantly, De Benedictis and Tajoli [105] studied the temporal evolution of the trade network, finding a strong and increasing heterogeneity of the countries' number of partners. Squartini et al. [102] found that most of the structural properties of the binary country–country network can be traced back to its degree sequence, whereas the same does not hold for the weighted representation where higher-order structural patterns cannot be explained by the countries' total export volumes [103]. König et al. [25] found that the binary country–country network exhibits a significantly nested structure, and proposed a network formation model to explain its topology [25,36] (see Section 5.3.1). Beyond the widely-studied World Trade network, other unipartite trade networks have been considered in the literature. For instance, some scholars [25,106] have investigated the topology of the global arm trade network, finding that it exhibits a significantly nested structure.

Besides aggregate trade networks, there is a growing interest in the topology of the trade networks for specific products [107] and industries [108]. Recent studies suggest that the trade networks of more sophisticated products are typically more centralized [109] and nested [107] than those of less complex products. While positive, the correlation between product complexity (as determined by node importance metrics, see Section 3.3.2) and trade-network centralization is far from being perfect [107,109], which suggests that complexity and nestedness may provide complementary information on a product.

Bipartite trade networks. Another way to look at World Trade data is to study the bipartite country–product network where countries are connected with the products they export. While the input data for this analysis are typically weighted, scholars have introduced procedures to “binarize” the network: for example, one can add a binary link from country i to product α if and only if the relative share of the product in the export basket of country i is larger than the World's average share [110]. Scholars have investigated the nested structure of the country–product network [29,52,111,112] (see Section 4.1.1), proposed generative mechanisms to explain its emergence [51,52] (see Section 4.3.4–5.3.4), and investigated its dependence on data regularization procedures and product aggregation schemes [113]. Besides, a recent stream of works have introduced and studied ranking algorithms for countries and products that enhance the nestedness of the country–product matrix [24,114–117] (see Sections 3.3.2 and 6.1.2).

The nested structure of the country–product network has deep implications for economic theories. Back in the XIX century, the comparative advantage theory by Ricardo [118] predicted that countries benefit from specializing on the products on which they have a comparative advantage. Besides, based on the found U-shape relationship between metrics for country diversification and income, Imbs and Wacziarg [119] claimed that “poor countries tend to diversify, and it is not until they have grown to relatively high levels of per capita income that incentives to specialize take over as the dominant economic force.” On the other hand, the empirical organization of the country–product network reported by recent literature [24,29] (see Fig. 3) indicates that, in fact, the most developed countries are highly diversified: they exhibit diverse export baskets by having a revealed comparative advantage [120] over a large number of products. By contrast, developing countries are only competitive in the export of products that are also exported by highly-diversified countries. Recent predictive schemes [45,121,122] based on network-based ranking algorithms suggest that country diversification is strongly correlated with the “economic fitness” of a country, and a country whose economic fitness is larger than that of countries with a similar level of development tend to economically grow in the future (see Section 4.3.3).

2.3.2. Contractor networks

In many economic systems, production is outsourced, which means that a firm “outsources” a part of its internal activity to an external company [123]. Scholars have identified both benefits and negative effects of outsourcing; we refer the interested reader to [123] for a review. Importantly for the present article, one can build a network that connects the firms (“manufacturers” or “designers”) that design and market products to the firms (“contractors”) that perform the activities that are necessary for the actual fabrication of the product [27,124].

Uzzi [124] analyzed the network of resource exchange between New York apparel firms over almost two years (1990–1991). He was interested in determining whether the structure of the transaction network can predict firms' failure. He found that (1) firms whose transaction volume is more heterogeneously distributed across the other firms (i.e., more generalist firms) are less likely to fail than firms that tend to only interact with few other firms (i.e., more specialized firms); (2) The topology of the partners' neighborhood heavily affects the probability that a firm fails: firms that are less likely to fail have neither too specialist nor too generalist partners.

By analyzing a 15-year designer–contractor dataset in the New York garment industry, a later study [28] revealed that the probability of firm failure is strongly affected by the firm's contribution to the overall nestedness of the manufacturer–contractor network (see Section 4.3.1). Besides, Saavedra et al. [27] found that manufacturer–contractor networks exhibit a significant nested structure, and proposed a parsimonious model with bipartite cooperation to explain the emergence of these structures (see Section 5.2.2).

Beyond designer–manufacturer networks, one can study other systems of contractors that involve a buyer and a seller. For example, Hernández et al. [125] analyzed 18-month dataset from the Boulogne-sur-Mer fish market. This fish market has the unique feature that the fish is sold through an auction directly from the buyers to the sellers, and transactions are recorded daily. Intriguingly, the aggregate buyer–seller network exhibits a significantly nested architecture, which excludes the emergence of blocks that would correspond to “niche” markets [125].

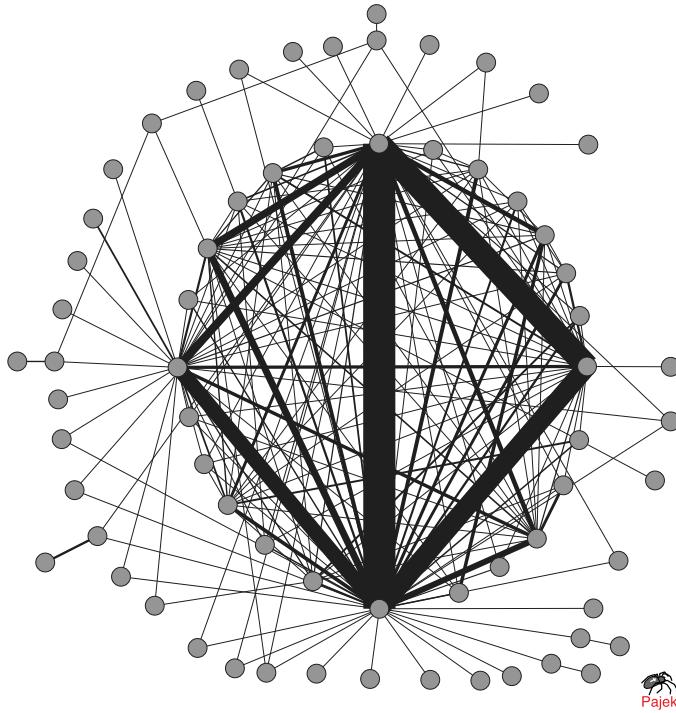


Fig. 4. The topology of the US interbank payment network (first quarter of 2004). The figure only shows the undirected links that represent the 75% of the total transaction flows. The network exhibits a clear core-periphery structure, with a core of densely interconnected nodes, and a periphery of nodes that are only connected with the core nodes. This topology can be considered as a special case of a nested topology (see Section 4.1.3). Source: Reprinted from [31].

2.3.3. Interbank networks

Understanding the topology of interbank networks is critical to assess their robustness, resilience, and effectiveness. Soramäki et al. [31] built an interbank transaction network based on the Fedwire Funds Service, a real-time gross settlement service. The topology of the network is shown in Fig. 4: it exhibits a clear distinction between a core of densely interconnected nodes, and a periphery of low-degree nodes that are only connected with nodes that belong to the core. Besides, the network is disassortative, meaning that high-degree nodes tend to be connected with lower-degree nodes, on average, than low-degree nodes (see Section 4.1.2). The network, therefore, exhibits two properties – core-periphery structure and disassortativity – that are strongly correlated with nestedness (see Sections 4.1.2–4.1.3); similar structural properties have been found in diverse interbank transaction and trade networks by subsequent studies (see [126–129], for instance).

Importantly, the core-periphery structure of interbank networks reveals important insights about financial crises: for example, Fricke and Lux [128] showed that the decline of interbank lendings during the 2008 financial crisis was mostly due to the core banks reducing their number of active outgoing links; Kojaku et al. [130] found that in correspondence of the 2008 crisis, the Italian interbank network transitioned from a structure with multiple core-periphery structures (see Section 7.3.3) to a bipartite structure [129] where transactions are mostly established between the core and periphery banks. The core-periphery structure can be considered as a particular case of nestedness (see Section 4.1.3), which suggests that interbank networks exhibit a significant level of nestedness as well. This hypothesis was recently corroborated by König et al. [25] who found that both an Austrian interbank network and the global banking network are significantly nested.

2.3.4. Spatial networks

In a similar way as spatial ecological networks connect spatial regions with the species that inhabit them, spatial economic networks connect spatial regions to the economic activities that are developed in them [29,30]. Bustos et al. [29] built a Chilean municipality-industry network by considering an industry i as present in a given municipality α if at least one firm classified in industrial classification i declared municipality α as its tax residence. In this way, approximately 700 different industries were connected with Chile's 347 municipalities [29]. They found that the resulting bipartite network exhibits a significantly nested structure, and exploited this property for link prediction, i.e., to predict the appearance and disappearance of industries in municipalities (see Section 4.3.2).

In a similar spirit, Garas et al. [30] built a city-firm network where 21 economic activities were connected with 1,169 cities in the world; also this network turned out to be significantly nested. From a node-level perspective, Gao

and Zhou [131] analyzed a bipartite network connecting industries with the 31 Chinese provinces, finding that ranking algorithms that seek to maximize the adjacency matrix nestedness (like the fitness–complexity algorithm [24]) provide a node score that is highly correlated with province-level macroeconomic indicators.

2.3.5. Communication networks in organizations

In social science, the idea that different topologies of communication networks lead to different performances of groups of individuals is not recent. Already in 1941, Leavitt [132] tested the performance of groups of individuals arranged in different communication topologies in solving puzzles. He found that a network with a star-like structure ("wheel structure" in [132]) performed better than more decentralized networks: the presence of a "generalist" who could be quickly reached by the other nodes seemed to speed up the problem-solving process. As reminded by Borgatti et al. [133], later studies have shown that as the complexity of the puzzle increases, decentralized networks tend to perform better.

This idea has motivated researchers to investigate how the structure of organizations impacts their performance and ultimate success. Importantly, organizations are made by the individuals who coordinate and constantly enhance internal knowledge. Communication networks (with organization members being nodes and communication events, as a proxy for information exchange, being links) play an important role in the structure of organizations. This fact has originated in the implicit knowledge that resides in the organization, which becomes uncovered by the patterns of communication [134].

Kogut and Zander [134] suggested that economic firms are "social communities in which individual and social expertise is transformed into economically useful products and services by the application of a set of higher-order organizing principles". Thus, firms exist because of a social community structured "by organizing principles that are not reducible to individuals" [134]. Firms are better in sharing and transferring knowledge among individuals than markets due to a fundamental dilemma: while higher specialization increases productivity due to the division of labor, it also increases the costs of communication and coordination among the agents [135]. Extant research has shown that the formal and informal networks of communications vastly differ [136,137]. The topology of this network, therefore, plays a role in determining the global properties of the organization.

Empirical work has found that intra-firm informal communication networks exhibit properties compatible with nestedness [43,138]. This is compatible with theoretical arguments mimicking how these networks evolve [36]. From a managerial perspective, an important question for future research is to assess whether nested communication flows are beneficial or harmful to the organization's efficiency.

2.3.6. Online communication networks and software development

Online social networks and social media platforms offer us an unprecedented amount of data on human activity. Often, these data come with fine-grained temporal information, which allows us to study how network topology varies over time. Valverde and Solé [139] have analyzed the network structure of Open-Source software development communities. They found that the system exhibits a strongly hierarchical structure where "an elite of highly connected and mutually communicating programmers control the flow of information generated by the OS community" [139]. Besides, the network is disassortative, which means that the average degree of the low-degree nodes' neighbors tends to be larger than that of the high-degree nodes' neighbors. Both properties are compatible with nestedness; these results suggest that communication networks between online community members might be a future field of application for nestedness analysis.

A different relevant network for analyzing software development is the network of dependencies and conflicts between software packages in a given operating system. The temporal evolution of the modularity of this network (see section 4.1.4 for the definition of modularity) has been also investigated by Fortuna et al. [140] who found that for the Debian GNU/Linux operating system, both modularity and the number of modules increased over time.

In the context of social-media analysis, Borge-Holthoefer et al. [32] studied the temporal evolution of nestedness and modularity for a bipartite user-hashtag network collected from Twitter. They created the bipartite network by collecting the memes related to the 2011 civil protests in Spain within a two-month time window (see [32] for the details). They found that a sharp transition from a modular to a nested topology occurs in the vicinity of a critical event – see Section 4.1.4 for details.

More recently, the temporal evolution of topological properties of Twitter communication networks has been investigated by Bastos et al. [141]. They focused on the centralization of the user–user communication network (relative to a specific topic, agriculture) in relation to the different level of specialization of shared information (as determined by unsupervised text classification techniques). Intriguingly, they found that a more centralized topology emerges when discussions become more technical, whereas generic discussions unfold over decentralized topologies. Their results confirm the premises of classical diffusion theories (see Chapter 9 in [142]) that posit that more centralized topologies are more appropriate for the dissemination of innovations that involve a high level of technical expertise.

2.4. Other classes of networks

Beyond ecological and socio-economic systems, other classes of networks have been found to exhibit nestedness. Kamilar and Atkinson [143] recently analyzed various datasets on the distribution patterns of cultural traits across populations of humans, chimpanzees, and orangutans. They found that for humans and chimpanzees, these patterns are

significantly nested, meaning that “cultures with a small repertoire of traits tend to comprise a proper subset of those traits present in more complex cultures” [143]. The same does not hold for orangutans, and we refer to the Discussion section in [143] for a detailed discussion of the possible reasons behind this discrepancy.

Cantor et al. [144] recently brought nestedness analysis to unipartite biological networks from six different levels of biological organization representing gene and protein interactions, complex phenotypes, animal societies, metapopulations, food webs and vertebrate metacommunities [144]. They found that nestedness emerges at various biological scales, and emphasized the importance of understanding the basic mechanisms behind its emergence. The core-periphery structure has received much attention in brain networks [145] where it has been found that the separation between temporal core and periphery changes can be used to predict individual differences in learning success. As we shall see in Section 4.1.3, a core-periphery topology is a special case of a nested topology.

Johnson et al. [146] analyzed various unipartite and bipartite networks of different types, including food webs, brain networks, metabolic networks, transportation networks and some popular network datasets (like the Zachary’s Karate Club network [147] and a collaboration network of Jazz musicians [148]). They found that a number of them exhibit significant nestedness — their detailed results can be found in the appendix S5 of their article [149].

In principle, any unipartite or bipartite network can be analyzed by means of nestedness metrics and null models. While, so far, the emergence and implications of nestedness have been mostly investigated in ecological and, more recently, socio-economic networks, we envision that they may gain additional interest from other research areas.

3. Observing nestedness: Metrics and null models

While perfectly nested networks can be unambiguously defined, they are rarely found in nature. Yet, many real networks exhibit highly-ordered structures where many pairs of nodes respect the definition of nestedness. This means that in such networks, given two nodes i and j such that $k_i < k_j$, most of node i ’s neighbors are also neighbors of node j . In these highly but imperfectly nested structures, the way we measure the level of nestedness can heavily affect our conclusions on the significance of the pattern and its relevance. This leads to several questions: how to best measure the degree of nestedness in real networks that are not perfectly nested? How to compare the level of nestedness across datasets of different size, density, and degree distribution? What is the impact of more basic network properties (e.g., the degree distribution) on the observed levels of nestedness?

This Section tackles these questions by providing an overview of the metrics (Section 3.1) and null models (Section 3.2) that have been introduced with the explicit goal to quantify the level of nestedness of a given system. Starting with the unexpected absences metric by Patterson and Atmar [21], we will survey various nestedness metrics including the popular nestedness temperature [22,70], NODF [150], and spectral radius [83]. Most of these metrics were originally introduced in the ecological literature, yet their realm of application has extended to economic and social networks.

In relation to the null models, there has been an intense debate in the ecological community over which null model should be preferred to infer the significance of the observed values of nestedness metrics [48,61,62]. We will cover the main strength and weaknesses of the available null models. The discussion will be further deepened in Section 4.1.1, where we will discuss the statistical relation between the degree distribution and nestedness.

Besides, as we have pointed out in the Introduction, one can also re-arrange the rows and columns of a given networks in such a way to minimize the number of violations of the nestedness condition. Such reshuffling of rows and columns can be interpreted as a node ranking algorithm, and it has important implications for the identification of vulnerable species in ecological systems [71], of competitive countries in international trade [24], and for the effectiveness of targeted attacks on the network⁶ [71,72]. The results of a quantitative comparison of existing ranking algorithms for bipartite networks [73] are presented in Section 3.3.

3.1. Metrics to quantify nestedness

In line with Ulrich [48], we classify metrics for nestedness into four main categories: gap-counting metrics (Section 3.1.1), overlap metrics (Section 3.1.3), distance metrics (Section 3.1.2), and eigenvalue-based metrics (Section 3.1.4). In principle, all these metrics apply to any binary matrix. For the sake of generality, unless otherwise stated, the definitions provided below refer to bipartite binary networks. The corresponding definitions for unipartite binary networks can be readily obtained by identifying the row-nodes with the column-nodes. Not all the metrics below have been generalized to weighted networks, yet we will provide some instances of metrics adopted in weighted networks.⁷

Before defining the metrics, an important caveat is that each of them can depend on basic network properties such as network size, density, and degree distribution; in order to compare nestedness metrics across networks of different size and density, performing a statistical analysis with a null model is a necessary step — this fundamental aspect is deepened in Section 3.2. To prevent redundancy, we will avoid repeating this caveat for each of the metrics defined below, yet the conscious reader needs to always keep it in mind.

⁶ The role of ranking algorithm for the identification of structurally important and vulnerable nodes is discussed in Section 6.1.2.

⁷ W-NODF [151] and spectral radius [83]: see Sections 3.1.3 and 3.1.4, respectively.

1	1	1	1	1	1	1	0
1	1	1	1	1	0	0	1
1	1	1	0	1	1	1	0
1	1	1	1	0	0	0	1
0	1	1	1	1	1	0	0
1	0	1	1	0	0	0	0
1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0

1	1	1	1	1	1	1	0
1	1	1	1	1	0	0	1
1	1	1	0	1	1	1	0
1	1	1	1	0	0	0	1
0	1	1	1	1	1	1	0
1	0	1	1	1	1	0	0
1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0

Fig. 5. An example of calculation of \mathcal{N}_0 (left) and \mathcal{N}_1 (right panel). Rows and columns are ordered by degree. *Left*, \mathcal{N}_0 . We scan each row from right to left. For each row i , we consider the filled spot (1) that corresponds to the column $\beta_{\min}(i)$ with the smallest degree; each zero at the left of such column, if it corresponds to a column α with degree $k_\alpha > k_{\beta_{\min}(i)}$, contributes to \mathcal{N}_0 . *Right*, \mathcal{N}_1 . We scan each row from left to right. For each row i , we consider the empty spot (0) that corresponds to the column $\beta_{\max}(i)$ with the largest degree; each filled element (1) at the right of such column, if it corresponds to a column α with degree $k_{\beta_{\max}(i)} > k_\alpha$, contributes to \mathcal{N}_1 .

3.1.1. Gap-counting metrics

Gap-counting metrics build on the observation that the empty and filled regions of a perfectly nested matrix are perfectly separated, as it is evident from the illustration in Fig. 1. This implies that there are no "absences" (i.e., empty spots) in the filled region, and no "presences" (i.e., filled spots) in the empty region. One can, therefore, evaluate the degree of nestedness of a given network by counting the number of violations of this property.

Unexpected absences and presences. The number of unexpected absences \mathcal{N}_0 introduced by Patterson and Atmar [21] counts how many times a node i is not a neighbor of a node α that has larger degree than its lowest-degree neighbor $\beta_{\min}(i)$ (i.e., $k_{\beta_{\min}(i)} := \min_{\alpha: A_{i\alpha}=1} \{k_\alpha\}$). In formulae,

$$\mathcal{N}_0 = \sum_i \sum_\alpha (1 - A_{i\alpha}) \Theta(k_\alpha - k_{\beta_{\min}(i)}). \quad (1)$$

In this formula, the sum over α is restricted to the pairs such that $k_\alpha > k_{\beta_{\min}(i)}$ through the Heaviside function $\Theta(k_\alpha - k_{\beta_{\min}(i)})$, where $\Theta(x) = 1$ if $x > 0$, $\Theta(x) = 0$ if $x \leq 0$. For each contribution to the sum, the $(1 - A_{i\alpha})$ factor excludes node i 's neighbors. Importantly, Patterson and Atmar noticed that their nestedness metric is a sum over the contributions from all the individual nodes, which implies that one can compare the contributions to nestedness from different nodes. A graphical illustration of the \mathcal{N}_0 metric can be found in Fig. 5.

A complementary perspective is offered by the number \mathcal{N}_1 of unexpected presences [152] metric which counts how many times a node i is neighbor of a node α that has smaller degree than its largest-degree non-neighbor $\beta_{\max}(i)$ (i.e., $k_{\beta_{\max}(i)} := \max_{\alpha: A_{i\alpha}=1} \{k_\alpha\}$). In formulae,

$$\mathcal{N}_1 = \sum_i \sum_\alpha A_{i\alpha} \Theta(k_{\beta_{\max}(i)} - k_\alpha). \quad (2)$$

A graphical illustration of the \mathcal{N}_1 metric can be found in Fig. 5.

One can also combine unexpected absences and presences into a single nestedness metric. This was done by Cutler [152], who considered one single node i of intermediate degree. In line with the definition of nestedness, absences of node i 's neighbors from the neighborhood of nodes with larger degree are counted as unexpected absences (U_i^A), whereas occurrences of node i 's non-neighbors in the neighborhoods of smaller-degree nodes are counted as unexpected presences (U_i^P). One chooses the node i^* that minimizes the sum of unexpected absences and presences $U_i^T = U_i^A + U_i^P$; accordingly, the degree of nestedness is defined as $U := \min_i U_i^T = U_{i^*}^T$. The U metric can be interpreted as the minimum number of steps to convert (either by filling empty spots or by deleting filled spots) the matrix in hand into a perfectly nested matrix.

Number of departures. For an ordered matrix, the number of departures, D , is defined as the number of times the absence of a node is followed by its presence in the neighborhood of the next lower-degree node [153]. Simply, the number of departures is the number of times the i th row does not interact with column α but the $(i+1)$ th row ($k_{i+1} < k_i$) does. Lomolino [153] estimated the statistical significance of the number of departures by comparing the measured D values with those obtained for randomly-ordered matrices. In the randomization, species distributions are unaltered, but islands are randomly ordered with respect to isolation and area. He calculated the normalized number of departures of each ordered matrix as:

$$\mathcal{N} = 100 \frac{(R - D)}{R}, \quad (3)$$

where D is the number of departures in the ordered matrix and R is the mean number of departures for randomized networks.

1	1	1	1	1	1	1	1	0
1	1	1	1	1	0	0	1	
1	1	1	0	1	1	1	0	
1	1	1	1	0	0	0	1	
0	1	1	1	1	1	0	0	
1	0	1	1	0	0	0	0	
1	1	0	0	0	0	0	0	
1	0	0	0	0	0	0	0	

1	1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	0	0
1	1	1	1	1	1	1	0	0
1	1	1	1	1	1	0	0	0
1	1	1	1	1	1	0	0	0
1	1	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0

Fig. 6. An example of calculation of the number of discrepancies, \mathcal{D} [154]. From the original matrix (left panel), one constructs a corresponding perfectly nested matrix (right panel) by shifting, for each row, all its filled elements to the right. The elements of the observed matrix that do not match the corresponding ones for the perfectly nested matrix are interpreted as “discrepancies” (marked in red in the left panel).

Deviations from a perfectly nested matrix. A perfectly nested matrix is one where the filled elements in each row are found as far to the left as possible, and the filled elements in each column are found as far to the top as possible. Motivated by such perfectly nested structure, Brualdi and Sanderson [154] construct, for a given adjacency matrix A , a perfectly nested matrix P by shifting the filled elements of A as far to the left as possible, while keeping their row fixed. Such shifting procedure generates matrices where each row-node i has the same degree k_i as the original matrix P , but the first k_i elements of row i from the left are filled. The discrepancy of a given matrix A can then be computed by summing over the rows the unexpected absences in each row i , i.e., the number of zeros observed in the first k_i elements from the left. In formulas,

$$\mathcal{D} = \sum_i \sum_{\alpha=1}^{k_i} (1 - A_{i\alpha}). \quad (4)$$

An illustration of the metric is provided in Fig. 6. Brualdi and Sanderson [154] assessed the statistical significance of the metric by comparing its observed value with its mean value in randomized networks generated with a null model that preserves exactly the row-nodes’ and column-nodes’ degree (Fixed–Fixed model, see Section 3.2.1).

3.1.2. Distance-based metrics

Distance-based metrics are based on a three-step computation. First, one determines the ideal line that separates the empty and filled regions of the matrix in a perfectly nested network. Second, one uses a suitable node-level ranking algorithm to rearrange the rows and columns of the adjacency matrix in such a way that its nestedness is maximized. Finally, for the rearranged matrix, one computes the distance of the unexpected elements (empty elements in the filled region, and filled elements in the empty region) from the ideal line. The nestedness temperature of the network is given by the sum of the contributions from all the unexpected elements of the rearranged matrix: highly nested networks are characterized by low temperature (see Eq. (5) below). Note that these metrics penalize more heavily unexpected absences and presences that are far away from the ideal line. Such assumption was originally motivated by arguments based on theories of island biogeography [22], as discussed below.

Interpretation in terms of fragmented habitats formation. Distance-based metrics are motivated by considerations based on species distribution patterns within naturally fragmented habitats [22]. Consider an original biota that subsequently fragments into a collection of islands due to natural causes. On each resulting island, there will be some species at larger risk of becoming extinct. Some species will indeed become extinct, and larger islands will be left, typically, with more species than smaller islands. The resulting system can be represented as a species-island bipartite network where each species is connected with the islands where its individuals are found.

If the species extinction order is exactly the same in each island, the resulting species-island network would be perfectly nested: each smaller island would only include a subset of the species that are found in larger islands. In this sense, nestedness is a property that signals the existence of *order* in the extinction patterns, whereas deviations from a perfectly nested structure can be interpreted as “statistical noise” [22]. Therefore, Atmar and Patterson made an explicit connection to statistical physics [155] and information theory [156]: similarly as a high-energy configuration of a system with many particles, the unexpected presence or absence of a species from an island is a “surprising event”.

Nestedness temperature and Nestedness Temperature Calculator (NTC). According to Ulrich [48], the matrix nestedness temperature T [22,70] is “by far the most popular metric for quantifying nestedness”. Matrix temperature depends on both the determination of the line of perfect nestedness and the matrix packing algorithm (details below), and different algorithms have been proposed for the temperature computation. The first algorithm proposed is the Nestedness Temperature Calculator (NTC) [22] which involves the following three steps:

1. **Determination of the line of perfect nestedness.** For a matrix with connectance $\rho = 0.5$, it is straightforward to identify the diagonal that bisects the matrix as the line of perfect nestedness (see Fig. 2 in [22]). For a matrix of connectance different from 0.5, the NTC software constructs the line of perfect nestedness geometrically (see Fig. 4 in [22]), whereas the BINMATNEST algorithm uses a function that determines the shape of the perfectly nested matrix (see Eq. (6) below).

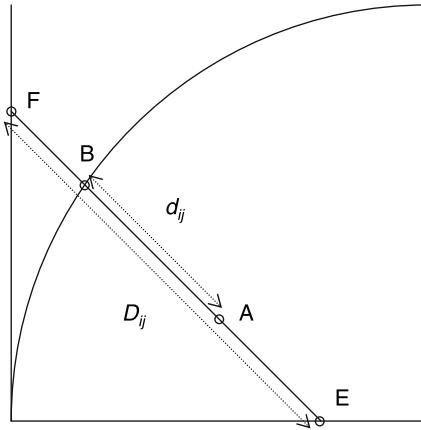


Fig. 7. A graphical illustration of the terms d and D needed in the temperature calculation [22,70]. The rows' and columns' labels are rescaled in such a way that the labels range from zero to one. From an unexpected presence or absence (in the figure, presence A below the line of perfect nestedness), one draws a straight line of slope -1 . The distance d_{ij} of the unexpected element A is then normalized by D_{ij} ; the nestedness temperature defined by Eq. (5) is given by all the unexpected elements' contributions $(d_{ij}/D_{ij})^2$.

Source: Reprinted from [70].

2. Packing the matrix. Atmar and Patterson [22] argued that “for a matrix's temperature to be calculated, a matrix must first be packed to a state of minimum unexpectedness”, and “row and column totals cannot be used as a reliable guide for packing”. By arranging the elements of a given matrix before computing the nestedness metric, we are not computing the nestedness associated with the degree or a specific property, but the maximum possible degree of nestedness in the system. We discuss the implications of this aspect and possible methods for packing the matrix in Section 3.3.

3. Computing the temperature. All the unexpected presences and absences contribute to the matrix temperature. The unexpectedness $u_{i\alpha}$ of a single (present or absent) link (i, α) is given by $u_{i\alpha} = (d_{i\alpha}/D_{i\alpha})^2$, where $d_{i\alpha}$ denotes the distance of (i, α) from the line of perfect nestedness (see Fig. 7), and $D_{i\alpha}$ is geometrically defined as explained in Fig. 7. The total unexpectedness U of the matrix is defined as

$$U = \frac{1}{NM} \sum_{(i,\alpha) \in \mathcal{U}} \left(\frac{d_{i\alpha}}{D_{i\alpha}} \right)^2, \quad (5)$$

where the sum is restricted to the set \mathcal{U} of unexpected element (i.e., empty elements above and filled elements below the line of perfect nestedness). The matrix temperature T is defined as $T = 100 U/U_{max}$, where $U_{max} = 0.04145$. A perfectly nested matrix has zero temperature, whereas a maximally non-nested matrix has a temperature equal to 100.

Variants of the NTC differs from the NTC in one or more of the three elements above: line of perfect nestedness determination, packing algorithm, and temperature computation. Besides, scholars have been also interested in developing software for a fast computation of the nestedness temperature over a large number of matrices [157]. Below, we briefly introduce two variants of the NTC: BINMATNEST and the τ -temperature.

BINMATNEST. The BINMATNEST algorithm [70] differs from the NTC in both the isocline determination procedure and in the packing algorithm. In particular, Rodríguez-Gironés and Santamaría [70] pointed out the following limitations of the NTC: the line of perfect nestedness is not uniquely defined, and the packing of the matrix is not optimal. In particular, as for the line definition problem, they pointed out that the definition implemented by the NTC does not identify a single line, but a set of curves [70]. To overcome this limitation, they defined the following function:

$$f(x; p) = \frac{0.5}{N} + \frac{N-1}{N} \left(1 - \left(1 - \frac{Mx - 0.5}{M-1} \right) \right). \quad (6)$$

The function depends on a parameter, p , that can be tuned to match the desired level of connectance ρ ; the resulting line of perfect nestedness obtained in real networks is essentially indistinguishable from those obtained by the NTC (see Fig. 2 in [70]). The packing algorithm implemented by the BINMATNEST algorithm and its relation with that implemented by the NTC is discussed in Section 3.3.

τ -temperature. The τ -temperature [158] differs from the NTC and the BINMATNEST algorithm in how the temperature is computed. Once the adjacency matrix is packed by a given algorithm, the τ -temperature of the matrix is proportional

to the Manhattan distance [158]

$$D = \sum_{i,\alpha} A_{i\alpha} (i + \alpha). \quad (7)$$

The τ -temperature is given by the ratio between D and the Manhattan distance observed in randomized networks that preserve the link density of the original network. Similarly as the Atmar and Patterson's temperature, the smaller the τ -temperature, the higher the matrix's degree of nestedness. The τ -temperature turns out to be positively correlated with the nestedness temperature T , yet such correlation is far from being perfect [158].

3.1.3. Overlap metrics

Overlap metrics are motivated by the definition of nestedness in terms of nodes neighborhoods – more specifically, by the property that in perfectly nested networks, the neighborhoods of nodes with lower degree is included in the neighborhoods of nodes with larger degree. Based on this consideration, one can then attempt to measure nestedness by evaluating how often neighbors of lower-degree nodes are also neighbors of larger-degree nodes.

Wright and Reeves' overlap-based metrics. In the context of species-island spatial networks, Wright and Reeves [159] built metrics for nestedness on the intuition that in a perfectly-nested network, if a species is found on an island, then it should be also found on richer islands. One can therefore quantify nestedness by aggregating the overlaps between the neighborhoods of pairs of islands. By introducing the overlap $O_{\alpha\beta} = \sum_i A_{i\alpha} A_{i\beta}$, they defined the N_c metric as [159]

$$N_c = \sum_{(\alpha,\beta)} O_{\alpha\beta} = \frac{1}{2} \sum_i k_i (k_i - 1) = \frac{N}{2} (\langle k_i^2 \rangle - \langle k_i \rangle), \quad (8)$$

where $O_{\alpha\beta} = \sum_i A_{i\alpha} A_{i\beta}$. Wright and Reeves [159] introduced an additional nestedness metric C by normalizing N_c as follows:

$$C = \frac{N_c - \mathbb{E}[N_c]}{\mathcal{N}_c^{(max)} - \mathbb{E}[N_c]}, \quad (9)$$

where $\mathbb{E}[N_c]$ is the expected value of N_c in randomized networks and $\mathcal{N}_c^{(max)}$ is the maximum value that can be attained by N_c . They computed $\mathbb{E}[N_c]$ based on a Equiprobable–Proportional null model (see Section 3.2.1) where the probability that a given species is found on a given island is proportional to the number of species in that island, obtaining $\mathbb{E}[O_{\alpha\beta}] = k_\alpha k_\beta / N$. Hence, one gets

$$\mathbb{E}[N_c] = \frac{1}{N} \sum_{(\alpha,\beta)} k_\alpha k_\beta = \frac{1}{2N} \sum_{\alpha,\beta} k_\alpha k_\beta = \frac{\langle k \rangle^2}{2N}. \quad (10)$$

It is also possible to compute the maximum value $\mathcal{N}_c^{(max)}$ which is achieved when $O_{\alpha\beta} = \min\{k_\alpha, k_\beta\}$. One thus obtains [159]

$$\mathcal{N}_c^{(max)} = \sum_{\alpha,\beta} k_\beta \Theta(k_\alpha - k_\beta) = \sum_\beta k_\beta (\beta - 1). \quad (11)$$

Wright and Reeves [159] observed that the C index "is free of strong dependence on the size of the presence–absence matrix, so that it is possible to compare the relative nestedness of different datasets". On the other hand, Brualdi and Sanderson [154] noticed that as the N_c metric is entirely determined by the row–column's degrees (see Eq. (8)), it is of no use if we wish to compare its observed value with its value in randomized networks that preserve exactly the network degree sequence (see also Section 4.1.1).

Nested Overlap and Decreasing Fill (NODF). Given a pair of row nodes (i, j) of the same class such that $k_i > k_j$, we expect their number of common neighbors $O_{ij} = \sum_\alpha A_{i\alpha} A_{j\alpha}$ to be equal to k_j for a perfectly nested network, and smaller than k_j a network that is not perfectly nested. One can then define the row-NODF N^R as [150]

$$N^R = \sum_{(i,j)} \frac{O_{ij}}{k_j} \Theta(k_i - k_j), \quad (12)$$

where Θ denotes the Heaviside function: $\Theta(x) = 1$ if $x > 0$, $\Theta(x) = 0$ if $x \leq 0$. This metric is maximal when $O_{ij} = k_j$ for all pairs (i, j) such that $k_j < k_i$. In the same way, one can define the column-NODF N^C as [150]

$$N^C = \sum_{(\alpha,\beta)} \frac{O_{\alpha\beta}}{k_\beta} \Theta(k_\alpha - k_\beta). \quad (13)$$

The degree of nestedness is quantified by the total NODF:

$$\eta = \frac{N^R + N^C}{\frac{N(N-1)}{2} + \frac{M(M-1)}{2}}. \quad (14)$$

S-NODF. We present here the metric introduced by Bastolla et al. [50] to facilitate the analytic computation of the impact of network structure on the co-existence of species in mutualistic systems. We refer to this metric as stable NODF (S-NODF) because it is a simple variant of NODF, yet it is more stable with respect to small perturbations of the data, as we shall discuss below.

The definition of rows' nestedness η^R reads [50]

$$\eta^R = \sum_{(i,j)} \frac{O_{ij}}{\min\{k_i, k_j\}}. \quad (15)$$

This definition⁸ differs from the one provided by NODF because it also includes the contribution of pairs of nodes with the same values of degree. In other words,

$$\eta^R = \mathcal{N}^R + \sum_{(i,j)} \frac{O_{ij}}{k_j} \delta(k_i, k_j), \quad (16)$$

where $\delta(k_i, k_j)$ denotes the usual Kronecker delta: $\delta(k_i, k_j) = 1$ if $k_i = k_j$, $\delta(k_i, k_j) = 0$ otherwise. Compared to the NODF contribution \mathcal{N}^R , η^R avoids penalizing pairs of rows with the same degree. The soundness of including pairs of rows with the same degree can be understood through the following example. Suppose that, in a mutualistic network, two pollinators i and j have a degree of 24 and 25, respectively, and suppose that they have 18 common neighbors (plants). This pair (i,j) contributes to \mathcal{N}^R with a contribution equal to $\mathcal{N}_{ij} = O_{ij}/k_i = 18/24 = 0.75$. However, suppose that a new interaction is observed between i and a plant that does not interact with j . In the updated network, both i and j have a degree equal to 25, and their contribution to NODF drops from 0.75 to zero: a small perturbation of the dataset can lead to a large variation in the pairwise contributions to NODF.

By contrast, in the example above, the contribution to η^R is only marginally affected by the pollinator i 's new interaction: when i 's degree increases from 24 to 25, the contribution $\eta_{ij} = O_{ij}/\min\{k_i, k_j\}$ to η^R decreases from $18/24 = 0.75$ to $18/25 = 0.72$, but it remains substantially larger than zero. For pairs of nodes of similar degree, the pairwise contributions to S-NODF are more stable with respect to small perturbations of node degree as compared to the pairwise contributions to the original NODF.

One can compute analogously the columns' nestedness; Bastolla et al. [50] defined the overall degree of nestedness η as:

$$\eta = \frac{\eta^R + \eta^C}{2}. \quad (17)$$

This nestedness index ranges from zero to one, where one corresponds to a perfectly nested network; it is highly correlated with NODF [50], yet it has the advantage of being more stable with respect to small structural perturbations. Differently from Eq. (17), one can normalize $\eta^R + \eta^C$ by the total number of pairs, as it was done for the original NODF (see Eq. (14)). This leads to the alternative definition:

$$\eta = \frac{\eta^R + \eta^C}{\frac{N(N-1)}{2} + \frac{M(M-1)}{2}}. \quad (18)$$

We argue that the definition of η provided by Eq. (18) should be preferred to that by Eq. (17), especially when the number of rows is substantially different than the number of columns.

JDM-NODF. Johnson et al. [146] considered a different normalization of nodes' pairwise overlap as compared to NODF and S-NODF. For a unipartite network, they defined the total nestedness as

$$\eta = \frac{1}{N^2} \sum_{i,j} \frac{O_{ij}}{k_i k_j}. \quad (19)$$

In a similar way as in [160], we refer to this metric as JDM-NODF after its inventors. This expression corrects a misprint of the metric introduced by Bastolla et al. [50] (see footnote 8 above), and it allows us to estimate the contribution η_i of each node to the global level of nestedness:

$$\eta_i = \frac{1}{N} \sum_j \frac{O_{ij}}{k_i k_j}. \quad (20)$$

W-NODF. The weighted NODF (W-NODF) [151] metric is specifically tailored to weighted networks. To assess the degree of nestedness of the columns of a weighted bipartite network, one compares the weight $w_{i\alpha}$ of each link (i, α) of a given column-node α with the weights $w_{i\beta}$ of the links between column-nodes of higher degree (β such that $k_\beta > k_\alpha$) and

⁸ Note that in the original paper [50] there was a typo such that the summation symbol appeared both in the numerator and in the denominator.

the same partner i . The columns' W-NODF $\mathcal{N}^{C,W}$ is larger if the weights $w_{i\alpha}$ tend to be smaller than the corresponding weights $w_{i\beta}$ for higher-degree column-nodes $k_\beta > k_\alpha$. The definition is the following [151]

$$\mathcal{N}^{C,W} = \sum_{(\alpha,\beta)} \frac{K_{\alpha\beta}}{k_\alpha} \Theta(k_\beta - k_\alpha), \quad (21)$$

where $K_{\alpha\beta}$ denotes the number of α 's weighted links with lower weights than the links between higher-degree column-nodes and the same row-node:

$$K_{\alpha\beta} = \Theta(k_\beta - k_\alpha) \sum_i A_{i\alpha} A_{i\beta} \Theta(w_{i\beta} - w_{i\alpha}). \quad (22)$$

One can define an analogous weighted NODF $\mathcal{N}^{R,W}$ for the rows, and then define the total weighted NODF η^W

$$\eta^W = \frac{\mathcal{N}^{R,W} + \mathcal{N}^{C,W}}{\frac{N(N-1)}{2} + \frac{M(M-1)}{2}}. \quad (23)$$

This equation is analogous to Eq. (14), with the difference that the rows' and columns' contributions depends not only on which pairs of nodes interact, but also on the weights of such interactions.

Normalizing overlap metrics based on null models. While one can in principle compare the observed values of any nestedness metrics with the values observed in randomized networks obtained with a given null model through standard statistical methods (like the z -score and p -value, see Section 3.2), scholars have also incorporated null-models effects directly into the overlap-based nestedness metrics. This has been done by including a (multiplicative or additive) normalization term which is typically determined by the expected values of the metrics under a suitable null model. The C metric by Wright and Reeves described above (see Eq. (9)) is an example of a normalized metric; we describe here analogous normalizations for overlap metrics.

Based on the JDM-NODF metric defined by Eq. (19), Johnson et al. [146] calculated the expected nestedness η_{CM} of a random unipartite network with expected degree sequence equal to the original network's degree sequence $\{k_1, \dots, k_N\}$, by posing $A_{ij} = k_i k_j / (N \langle k_i \rangle)$ in Eq. (19). We readily obtain $\eta_{CM} = \langle k^2 \rangle / (N \langle k \rangle^2)$; if one is interested in capturing the nestedness that cannot be explained by the degree sequence, it is therefore useful to consider the normalized JDM-NODF

$$\eta = \frac{\tilde{\eta}}{\eta_{CM}} = \frac{\langle k \rangle^2}{N \langle k^2 \rangle} \sum_{i,j} \frac{O_{ij}}{k_i k_j}. \quad (24)$$

This metric was used to reveal the relation between nestedness and degree-degree correlations in synthetic and real networks (see Section 4.1.2).

In a similar spirit but with a different methodology, Solé-Ribalta et al. [40] considered a variant of NODF that compares the observed level of nestedness (as determined by a variant of the NODF function defined by Eq. (14)) with the expected nestedness under a suitable null model. In formulas, their normalized nestedness metric is given by

$$\tilde{\mathcal{N}} = \frac{2}{N+M} \left\{ \sum_{i,j} \frac{O_{ij} - \langle O_{ij} \rangle}{(N-1)k_j} \Theta(k_i - k_j) + \sum_{\alpha,\beta} \frac{O_{\alpha\beta} - \langle O_{\alpha\beta} \rangle}{(M-1)k_\beta} \Theta(k_\alpha - k_\beta) \right\}. \quad (25)$$

Solé-Ribalta et al. [40] evaluated the expected row-row overlap $\langle O_{ij} \rangle$ and column-column overlap $O_{\alpha\beta}$ as $\langle O_{ij} \rangle = \langle O_{ji} \rangle = k_i k_j / M$ and $O_{\alpha\beta} = k_\alpha k_\beta / N$, respectively.⁹ The behavior of this function deviates from that of the original NODF for dense matrices where the expected pairwise overlaps with the largest-degree nodes tend to be large and, as a result, the respective contributions to $\tilde{\mathcal{N}}$ tend to be small [40].

3.1.4. Eigenvalue-based metrics

What can the eigenvalues of the network's adjacency matrix tell us about nestedness? The last metric for nestedness considered here is the *spectral radius* [83], which builds upon graph-theoretical results on the spectral properties of the adjacency matrix of bipartite networks [82,161]. The spectral radius $\rho(A)$ of a network is simply defined as the dominant eigenvalue of the network's adjacency matrix A [2]. Staniczenko et al. [83] considered $\rho(A)$ as a metric to quantify nestedness. This interpretation is motivated by the two following theorems:

- Consider all the connected bipartite networks composed of S nodes and E edges. The network with the largest spectral radius is a perfectly nested network. This theorem has been proved by Bell et al. [82].

⁹ The two expected values $\langle O_{ij} \rangle = \langle O_{ji} \rangle = k_i k_j / M$ is obtained by and $O_{\alpha\beta} = k_\alpha k_\beta / N$ are obtained with the PE and EP null model, respectively. In the PE (EP) model, the probability that a row-column pair of nodes interact is proportional to the degree of the row-node (column-node) – see Section 3.2.1 for details.

- Consider all the connected bipartite networks composed of N row-nodes, M column-nodes, and E edges. The network with the largest spectral radius is a perfectly nested network. This theorem has been proved by Bhattacharya et al. [161].

While the number of row-nodes and column-nodes is allowed to vary in the first theorem (provided that their sum is equal to S), the second theorem considers all networks with a fixed number of both row- and column-nodes. Therefore, if perfectly nested structures lead to the largest spectral radius for bipartite networks, the spectral radius $\rho(A)$ can be used to quantify the degree of nestedness of a given network: the larger $\rho(A)$, the more nested the network. The advantage of $\rho(A)$ is that it can be computed in short time through well-established techniques such as the power-method [11,162]. Besides, one can readily apply it to weighted networks — in that case, one simply needs to measure the spectral radius of the weighted adjacency matrix.

While maximal $\rho(A)$ is associated with a perfectly nested structure, it remains to be determined whether deviations from the maximal $\rho(A)$ should be deemed as statistically significant or not. To address this question, Staniczenko et al. [83] defined the p -value of a network under a given null model as the probability that a network randomized under that null model exhibits a larger $\rho(A)$. Such definition depends critically on the null model that is chosen for the randomization. For unweighted bipartite networks, Staniczenko et al. [83] considered three null models of different degree of conservativeness (both a model that preserves the degree sequence and two models that do not, see Section 3.2.1). For weighted bipartite networks, Staniczenko et al. [83] consider an additional null model that shuffles the weight values in the adjacency matrix A but not their position.

3.2. Including a null model

Given a network and a structural pattern of interest, one is typically interested in assessing its statistical significance: is the observed structural pattern compatible with that observed in randomized networks which preserve (a small number of) interesting macroscopic properties of the original network? To address this question for nestedness, scholars have introduced various null models and statistical tests, and applied them to the nestedness metrics introduced in Section 3.1.

How to choose the null model has been widely debated in the ecology literature. Back in the 70 s, among the first scholars who attempted to infer the significance of structural patterns in ecological networks, Sale [75] analyzed networks where species are connected with the resources they use, and he generated synthetic random networks by keeping fixed the number of resources used by each species, and allowing the number of species using a resource to vary (Fixed-Equiprobable model, see Section 3.2.1). Connor and Simberloff [65] used a null model that preserves exactly species' and islands' degrees in spatial networks, in the same spirit as the popular configuration model [163–165]. In 1982, Gilpin and Diamond [77] used a null model which preserves, on average, the species' and islands' degree; their model might be considered as a precursor of the popular Chung-Lu model [166].

A variety of additional randomization procedures have been introduced in the 80 s and 90 s in the ecological literature [48,60]; the basics of such models are provided in Section 3.2.1. More recently, Squartini and Garlaschelli [167] introduced a physics-rooted framework to compute analytically the expected properties of networks with a given expected degree sequence. In the same vein as the computation of thermodynamic properties in statistical mechanics [168], their calculation is based on constructing a maximum-entropy ensemble of networks with expected degree sequence equal to the original network's degree sequence (see Section 3.2.2). This statistical-physics approach has been exploited recently to assess the significance of nestedness in economic [112], and ecological [169] networks.

3.2.1. Nine basic classes of randomization procedures

When randomizing a given bipartite adjacency matrix, one has essentially two possible choices to make – one for the row-nodes and one for the column-nodes – about the preservation of the individual nodes' degree. In line with Gotelli [60], for both row-nodes and column-nodes, we consider three options: (1) preserving exactly the individual nodes' degree (the degree is *fixed*); (2) preserving on average the nodes' degree (the interaction probability is *proportional* to degree); (3) not preserving the nodes' degree, and assuming that all the nodes have the same probability to interact with the nodes of the other guild (the pairwise interactions are *equiprobable*). As the choice among such three options can be made independently for rows and columns, we have $3^2 = 9$ resulting classes of null models (see Table 4).

We emphasize that as soon as one has chosen the properties to keep fixed while randomizing the rest, there are multiple possible implementations for the resulting model. Surveying all the possible implementations falls out of the scope of this review, but we shall provide the main ideas of some of them. It is also worth noticing that scholars have started investigating more nuanced scenarios where one can “tune” the level of discrepancy of row-nodes' and column-nodes' degree (with respect to the observed degrees) in a continuous manner [64]. These recent developments are not included in this review; nevertheless, they demonstrate that research on null models is still active in ecology, and consensus on the answers to fundamental questions (e.g., which null model is more suitable for a given network and a given structural pattern?) has yet to be achieved.

Below, we provide details for the nine basic classes of null models of Table 4, we provide some of the alternative names that have been used in the literature, and we point to some of the papers that have used them.

Table 4

Nine basic classes of randomization procedures, corresponding to nine different null models. Each randomization procedure depends on two independent choices. The first choice is whether to preserve the row-nodes' degree. The three possible options are to not preserve it (Row Equiprobable), to preserve it exactly (Row Fixed), or to preserve it on average (Row Proportional). The second choice is whether to preserve the column-nodes' degree. Again, the three possible options are to not preserve it (Column Equiprobable), to preserve it exactly (Column Fixed), or to preserve it on average (Column Proportional). Each pair of choices leads to a specific class of randomization procedures (see the main text for a detailed explanation of each of them). This Table is inspired by Table 2 in [60].

	Column Equiprobable	Column Fixed	Column Proportional
Row Equiprobable	EE	EF	EP
Row Fixed	FE	FF	FP
Row Proportional	PE	PF	PP

- **Row Equiprobable–Column Equiprobable (EE) models.** All interactions are equiprobable, independently of the row-nodes' and column-nodes' degree. In a possible randomization procedure, the E original links are re-assigned to randomly selected pairs (i, α) of nodes; the probability that pair (i, α) is filled with a link is given by $1/(NM)$ [60]. Alternative names for this model are: SIM1, R00 [48,60], Type I model [23]. In ecology, scholars who have used this null model include: Atmar and Patterson [22], Gotelli [60]. In the graph-theoretical language, the EE model generates Erdős-Rényi graphs [10,11].
- **Row Equiprobable–Column Fixed (EF) models.** The column-nodes' degree sequence $\{k_\alpha\}$ is kept fixed, whereas their partners are chosen at random. In a possible randomization procedure, for each column α , the k_α original links are re-assigned to randomly selected row-nodes; the probability that node i is chosen as a partner is given by $1/N$ [60]. An alternative name for this model is SIM3 [60]. In ecology, scholars who have used this null model include Gotelli [60].
- **Row Equiprobable–Column Proportional (EP) models.** The probability that a pair of nodes interact is proportional to the column-node degree, whereas all row-nodes are equiprobable. In a possible randomization procedure, the E original links are re-assigned to randomly selected pairs (i, α) of nodes; the probability that pair (i, α) is filled with a link is given by $k_\alpha/(NE)$ [60]. An alternative name for this model is SIM6 [60]. In ecology, scholars who have used this null model include Gotelli [60].
- **Row Fixed–Column Equiprobable (FE) models.** The row-nodes' degree sequence $\{k_i\}$ is kept fixed, whereas their partners are chosen at random. In the randomization procedure, for each row i , the k_i original links are re-assigned to randomly selected column-nodes; the probability that node α is chosen as a partner is given by $1/M$ [60]. Alternative names for this model are: SIM2, R0, Random0 [60]. In ecology, scholars who have used this null model include: Sale [75], Inger and Colwell [76], Winemiller and Pianka [170], Patterson and Atmar [21], Gotelli [60].
- **Row Fixed–Column Fixed (FF) models.** The row-nodes' and column-nodes' degree are exactly preserved. Alternative names for this model are: SIM9 [60]. In ecology, scholars who have used this null model include: Connor and Simberloff [65], Diamond and Gilpin [171], Brualdi and Sanderson [154], Gotelli [60]. A popular randomization procedure¹⁰ to keep a network's degree sequence fixed is based on pairwise link swapping (see [165,172] and Fig. 8). In the network science language, the FF model can be interpreted as the most constrained version of the popular configuration model [11].
- **Row Fixed–Column Proportional (FP) models.** The row-nodes' degree sequence $\{k_i\}$ is kept fixed, whereas their partner column-nodes are chosen with probability proportional to their degree. In a possible randomization procedure, for each row i , the k_i original links are re-assigned column-nodes randomly selected with probability k_α/E [60]. Alternative names for this model are: SIM4 [60]. In ecology, scholars who have used this null model include: Coleman et al. [178], Graves and Gotelli [179], Gotelli [60].
- **Row Proportional–Column Equiprobable (PE) models.** The probability that a pair of nodes interact is proportional to the row-node degree, whereas all column-nodes are equiprobable. In a possible randomization procedure, the E original links are re-assigned to randomly selected pairs (i, α) of nodes; the probability that pair (i, α) is filled with a link is given by $k_i/(ME)$ [60]. An alternative name for this model is SIM7 [60]. In ecology, scholars who have used this null model include Gotelli [60].
- **Row Proportional–Column Fixed (PF) models.** The column-nodes' degree sequence $\{k_\alpha\}$ is kept fixed, whereas their partner row-nodes are chosen with probability proportional to their degree. In a possible randomization procedure, for each column α , the k_α original links are re-assigned to row-nodes randomly selected with probability given by k_i/E [60]. Alternative names for this model are: SIM5, R1, Random1 [60]. In ecology, scholars who have used this null model include: Abele and Patton [180], Connor and Simberloff [65], Patterson and Atmar [21], Gotelli [60].
- **Row Proportional–Column Proportional (PP) models.** The probability that a pair of nodes interact is proportional to both nodes' degrees. In a possible randomization procedure, the E original links are re-assigned to randomly selected pairs (i, α) of nodes; the probability that pair (i, α) is filled with a link is given by $k_i k_\alpha / E^2$ [60]. An alternative name for this

¹⁰ Alternative randomization procedures have been introduced in the literature. We refer the interested reader to [172–177] for various randomization procedures and their validation; to Appendix F of the book [9] for a general overview in ecology.

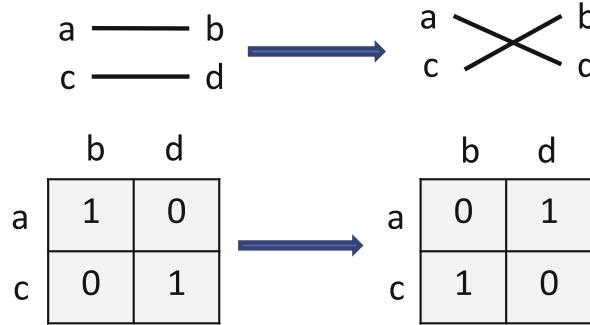


Fig. 8. An illustration of a popular implementation [165,173] of the Fixed-Fixed null model which randomizes the network by preserving exactly the nodes' degree. One randomly selects two links of the network (a, b) and (c, d); the two selected links are replaced by two new links, (a, d) and (b, c), if these two links did not exist before. In terms of the adjacency matrix, such swap corresponds to a swap of “checkerboard units” [182], where the initial and final configurations of the depicted 2×2 matrix are both checkerboard units.

class of models is SIM8 [60]. In Bascompte et al.'s implementation [23], the probability that a matrix element A_{ia} is filled is given by $(k_i + k_a)/2$ – the resulting model is referred to as Type II model [23]. In ecology, scholars who have used this class of null models include: Gilpin and Diamond [77], Gotelli [60], Bascompte et al. [23], among many others. Beyond ecology, randomization procedures that preserve, on average, the network's degree sequence are popular in the network science literature. Among the null models that belong to the PP family, we find the popular Chung–Lu model [166], the maximum-entropy models that are detailed in Section 3.2.2, generalized hypergeometric ensembles [181].

The careful reader may have noticed that interchanging the roles of row-nodes and column-nodes is equivalent to swapping pairs of classes of null models that fall out of Table 4's diagonal. In other words, if we transpose the original adjacency matrix of the network, the effect of EP models acting on A^T is equivalent to the effects of the corresponding PE models acting on A . For this reason, the PE and EP models can be seen as equivalent; nevertheless, in line with previous works [60,64], we preferred to present separately pairs of models that are equivalent upon transposition of the adjacency matrix, thereby implicitly assuming that the assignment of the nodes to A 's rows and columns is fixed.¹¹

Which null model to choose? When analyzing a collection of empirical networks, the choice of a null model critically affects the fraction of networks that result as “significantly nested”. Scholars have pointed out that while applying the EE model is the simplest way to randomize a given network, assuming equiprobable interactions might lead to highly unrealistic random networks. Ulrich [48] summarizes this discontent by pointing out that “there is growing acceptance that null models that do not consider species-specific differences and variability among sites should not be used in biogeographic studies [61,66,183,184] and even in analyses of interaction matrices [185–187]”. Based on their results, Joppa et al. [67] argue that “loose constraints invariably lead to the conclusion of significant nestedness”, and that such loosely-constrained null models “create consumers that differ in how generalized their diets are and this confounds the conclusion”. In other words, there is agreement that the EE model is sensitive to Type I errors: it can detect a significant nestedness in networks when there is none [62,160].

On the other hand, the FF model is the most conservative model, as it preserves exactly both the row-nodes' and the column-nodes' degree. Both the FF and the PP model set out to address the following natural question: are networks still significantly nested when one constrains the degree sequence to be the same (exactly or on average) as that of the original network? The use of a more conservative null model substantially mitigates the fraction of networks that are found as “significantly nested”. Yet, the FF model is more heavily affected by Type II errors [61]; we will come back to this point in Section 4.1.1.

Another potential issue with the randomization procedures described above (except for the FF model) is that they might generate “degenerate matrices” where some nodes have a degree equal to zero. To overcome this potential shortcoming, one can reject the degenerate networks (as done by the Swappable-Swappable model [83,160]), or perform the swapping between two matrix elements only if it does not cause one of the nodes' degree to drop to zero (as done by the Cored–Cored model [160]).

Finally, using randomized networks to generate random networks has two additional limitations: (1) it might be computationally slow when one attempts to preserve exactly the nodes' degree; (2) it relies on an arbitrary choice on

¹¹ For an ecological bipartite network where the assignment of the nodes to the A's rows and columns is fixed (e.g., a mutualistic network where A's rows represent insects and columns represent plants), there are ecological reasons to randomize differently the two groups of nodes. For example, in a plant–pollinator network, plants might be subject to more phenotypical constraints than pollinators, thereby restraining their degree of specialization. Instead, insect pollinators might exhibit a higher flexibility to change their diet. In a similar fashion, for spatial networks, varying the richness of a species or modifying its spatial distributions correspond to different ecological assumptions.

the number of independent realizations of the randomization procedure. We would not face those two problems if we were able to compute analytically the expected values of the structural properties of interest. In Section 3.2.2, we will see that this is made possible by a compelling analogy with the problem of computing the average value of macroscopic observables in statistical mechanics.

3.2.2. Computing the expected nestedness: Maximum-entropy approach

Null models based on network randomization have the disadvantage that the results based on them depend on the number of performed independent randomizations. Increasing the number of randomizations might make the results more robust, but substantially increase the computational time, especially for large networks. To overcome these limitations, a recent stream of literature [167,181] has provided theoretical frameworks to analytically compute the expected properties of random networks with a set of fixed macroscopic properties, for network properties that can be expressed in terms of the adjacency matrix A . In particular, Squartini and Garlaschelli [167] introduced a maximum-entropy framework to analytically calculate the expected value of network structural properties that can be expressed in terms of the network's adjacency matrix A .

In the maximum-entropy framework, one calculates the expected value of the network properties over a maximum-entropy ensemble of networks which preserves, on average, some network properties – the degree sequences, in all the cases considered here. This relieves us of running numerical simulations, which makes the null-model expectations faster to be computed and independent of the choice on the number of independent realizations of the adopted randomization procedure. Importantly, Squartini and Garlaschelli's framework has been also generalized to weighted [188,189] and bipartite networks [112]. The maximum-entropy method for unipartite networks has been applied to the country–country export network [102] and to the interbank network [190]. Their generalization to bipartite networks has been applied to assess the significance of nestedness in bipartite World Trade networks [112] and mutualistic networks [169] (see Section 4.1.1). We provide here the basic ideas behind the method, and we refer the interested reader to [112,167,169] for all the details.

Maximum-entropy approach in unipartite networks. In the maximum-entropy approach [167], we are interested in computing the expected value of a given observable property \mathcal{Q} of the network over a maximum-entropy ensemble of graphs with a given set of constrained properties. Let us denote the value of observable \mathcal{Q} in the observed network \mathcal{G}^* by \mathcal{Q}^* , the value of \mathcal{Q} in a generic network \mathcal{G} as $\mathcal{Q}(\mathcal{G})$. Let us denote by \mathcal{C} the vector of constrained properties of the maximum-entropy ensemble that we will construct. In the following, the constrained property is the complete network degree sequence; in this case, the vector \mathcal{C} has N elements, and $C_i = k_i^*$.

In Squartini and Garlaschelli's framework [167], one seeks to find a probability distribution $P(\mathcal{G})$ over the ensemble $\{\mathcal{G}\}$ of possible random graphs that maximizes the entropy

$$S[P(\mathcal{G})] = - \sum_{\mathcal{G}} P(\mathcal{G}) \log(P(\mathcal{G})) \quad (26)$$

under the constraints

$$\begin{aligned} \sum_{\mathcal{G}} P(\mathcal{G}) &= 1, \\ \sum_{\mathcal{G}} P(\mathcal{G}) k_i(\mathcal{G}) &= k_i^* \quad \text{for all } i \in \{1, \dots, N\}. \end{aligned} \quad (27)$$

Basically, we aim to maximize the entropy of the network probability distribution while keeping fixed the average nodes' degree.

The problem is analogous to the entropy maximization in statistical mechanics [168] where, given a system with a large number of microscopic constituents, one seeks to maximize the entropy of the probability distribution over the allowed microscopic configurations while keeping fixed the average total energy of the system. Therefore, we know from statistical mechanics that the solution of the problem is given by the canonical distribution

$$P(\mathcal{G}) = \frac{\exp(-H(\mathcal{G}, \boldsymbol{\theta}))}{Z(\boldsymbol{\theta})}, \quad (28)$$

where $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_N\}$ is a vector of N Lagrange multipliers, and

$$H(\mathcal{G}, \boldsymbol{\theta}) = \boldsymbol{\theta} \cdot \mathbf{k}(\mathcal{G}) = \sum_i \theta_i k_i(\mathcal{G}). \quad (29)$$

is the Hamiltonian of the system, and $Z(\boldsymbol{\theta}) := \sum_{\mathcal{G}'} \exp(-H(\mathcal{G}', \boldsymbol{\theta}))$ is the partition function. The optimal Lagrange multipliers $\hat{\boldsymbol{\theta}}$ can be found by first noticing that the expected value of A_{ij} in the maximum-entropy ensemble is given by

$$\langle A_{ij} \rangle = \frac{e^{-\theta_i - \theta_j}}{1 + e^{-\theta_i - \theta_j}}; \quad (30)$$

by enforcing the constraint $\langle \mathbf{k}(\mathcal{G}') \rangle = \mathbf{k}$, we obtain the equation:

$$\sum_{j \neq i} \frac{\exp(-\hat{\theta}_i - \hat{\theta}_j)}{1 + \exp(-\hat{\theta}_i - \hat{\theta}_j)} = k_i^* \quad \text{for all } i \in \{1, \dots, N\}. \quad (31)$$

Eq. (31) corresponds to a system of N equations that has a unique solution $\hat{\theta}$. We refer the interested reader to [167] for the full derivation of these equations.

Topological observables can be considered as functions $\Omega(A)$ of the network's adjacency matrix. By denoting as $A^{(exp)}$ the matrix whose element $A_{ij}^{(exp)} = \langle A_{ij} \rangle(\hat{\theta})$, we obtain the expected value of observable Ω in the maximum-entropy ensemble as $\Omega(A^{(exp)})$. One can also compute the standard deviation of observable Ω in the maximum-entropy ensemble as

$$\sigma[\Omega] \simeq \sqrt{\sum_{i,j} \left(\sigma^*[A_{ij}^{(exp)}] \frac{\partial \Omega[A]}{\partial A_{ij}} \Big|_{A_{ij}=A_{ij}^{(exp)}} \right)^2}, \quad (32)$$

where $\sigma^*[A_{ij}^{(exp)}] = \sqrt{A_{ij}^{(exp)}(1 - A_{ij}^{(exp)})}$ (see Eq. B31 in [167] and its derivation).

Maximum-entropy approach in bipartite networks. The framework for bipartite networks [112,169] is conceptually analogous. Differently from unipartite networks, one needs two sets θ^R and θ^C of Lagrange multipliers for row-nodes and column-nodes, respectively. The canonical distribution is given by

$$P(\mathcal{G}) = \frac{\exp(-H(\mathcal{G}, \theta^R, \theta^C))}{Z(\theta^R, \theta^C)}, \quad (33)$$

where $\theta^R = \{\theta_1^R, \dots, \theta_N^R\}$ is the vector of Lagrange multipliers for the N row-nodes, $\theta^C = \{\theta_1^C, \dots, \theta_M^C\}$ is the vector of Lagrange multipliers for the M column-nodes, Z is the normalization factor, and the Hamiltonian function is given by

$$H(\mathcal{G}, \theta^R, \theta^C) = \sum_i \theta_i^R k_i(\mathcal{G}') + \sum_\alpha \theta_\alpha^C k_\alpha(\mathcal{G}'), \quad (34)$$

Again, one obtains the expected values of the adjacency matrix elements as

$$A_{i\alpha}^{(exp)} = \frac{\exp(-\hat{\theta}_\beta^C - \hat{\theta}_i^R)}{1 + \exp(-\hat{\theta}_\beta^C - \hat{\theta}_i^R)}, \quad (35)$$

In the same way as for unipartite networks, the values of the Lagrange multipliers are determined by fixing the average degree of row-nodes and column-nodes, which leads to the following set of equations [169]

$$\begin{aligned} \sum_\beta \frac{\exp(-\hat{\theta}_\beta^C - \hat{\theta}_i^R)}{1 + \exp(-\hat{\theta}_\beta^C - \hat{\theta}_i^R)} &= k_i^* \quad \text{for all } i \in \{1, \dots, N\}, \\ \sum_j \frac{\exp(-\hat{\theta}_\alpha^C - \hat{\theta}_j^R)}{1 + \exp(-\hat{\theta}_\alpha^C - \hat{\theta}_j^R)} &= k_\alpha^* \quad \text{for all } \alpha \in \{1, \dots, M\}. \end{aligned} \quad (36)$$

and, consequently, the expected values of an observable $\Omega(A)$ as $\langle \Omega(A) \rangle = \Omega(A^{(exp)})$. The standard deviation of $\Omega(A)$ can be obtained with an equation analogous to Eq. (32) (see [169] for details).

Using the maximum-entropy framework to compute the expected nestedness of a network. The maximum-entropy analytic framework can be used to compute the expected value and standard deviation of any structural observable $\Omega(A)$ that explicitly depends on the network's adjacency matrix A . Some of the nestedness metrics introduced in Section 3 are explicit functions of A . A prominent example is NODF: the analytic calculation of the expected NODF for a set of mutualistic networks was carried out by Borràs et al. [169]. To perform the calculation, essentially, one replaces A with $A^{(exp)}$ as determined by Eq. (35) in the NODF definition (Eq. (14)). Importantly, Borràs et al. [169] were also able to compute the standard deviation of NODF, which allowed them to compute the z-score (see Section 3.2.3) of the empirically observed NODF values. We refer to Section 4.1.1 for a discussion of their results. Importantly, one can also use the maximum-entropy framework to compute “algorithmic” nestedness metrics (such as the nestedness temperature introduced in Section 3.1.2) by sampling links from the analytic probability of interaction between nodes (encoded in the $A^{(exp)}$ matrix), and then computing the moments of the algorithmic metrics over the set of generated sampled networks.

3.2.3. Statistical tests

The randomization procedures presented in Section 3.2.1 can be used to generate multiple independent randomized adjacency matrices, and to accumulate thereby statistics on the metrics for nestedness. Alternatively, one can compute

analytically the expected degree of nestedness based on a maximum-entropy ensemble of networks with fixed average degree distribution (Section 3.2.2). How to use this information to quantify the deviation of the degree of nestedness observed in a given real network from that expected in the corresponding random networks? Statistical tests allow us to address this question; we present here the most common statistical tests which are typically used in nestedness analysis.

In the following, we suppose that we are interested in one particular nestedness metric \mathcal{N} and a given observed network \mathcal{G}^* . We denote by \mathcal{N}^* the value observed in the real network of interest, and by $\mu(\mathcal{N})$ and $\sigma(\mathcal{N})$ the mean and the standard deviation, respectively, of the nestedness metric in the ensemble of random networks that are generated with the chosen null model. All the statistical tests below are based on the comparison between the observed level of nestedness \mathcal{N}^* and the statistical properties of \mathcal{N} over the ensemble of random networks.

z-Score. The *z-score* [48,160] is calculated as

$$z(\mathcal{N}) = \frac{\mathcal{N}^* - \mu(\mathcal{N})}{\sigma(\mathcal{N})}.$$

It tells us the distance from the mean of a given observed value of nestedness \mathcal{N}^* , in units of standard deviations. By randomizing empirical ecological networks with the FF model, Almeida-Neto et al. [150] found that $z(\mathcal{N})$ is only weakly correlated with network size and connectance for the overlap metric by Wright and Reeves (see Eq. (9)), the discrepancy index by Brualdi and Sanderson (see Section 3.1.1), the NODF metric, and matrix temperature (as determined by the NTC algorithm). The *z-scores* can be also used to compare the information provided by different metrics for nestedness (see Fig. 11 below and the related discussion in Section 4.1.1).

p-Value. The *p-value* [160] of an observed value of nestedness \mathcal{N}^* is the probability that a randomized matrix exhibits a degree of nestedness that is larger than \mathcal{N}^* . Low values of *p* ($p \rightarrow 0$) indicate that the input matrix is highly nested relative to the null distribution; usually, a threshold value λ (e.g., $\lambda = 0.05$ or 0.01) is used to denote a statistically significant level of nestedness ($p < \lambda$). For matrices where no randomized network is more nested than the input matrix, one can conservatively assign $p < R^{-1}$ where R is the number of independently generated random matrices [160].

3.3. Nestedness maximization: Packing the adjacency matrix

In the previous Section, we have introduced metrics to quantify the level of nestedness of a given network. The discussed metrics are mostly based on the definition of nestedness by degree: a pair of nodes respects the nestedness condition if the neighborhood of the node with lower degree is included in the neighborhood of the node with larger degree. This definition crucially depends on the ranking of the nodes by their degree.

On the other hand, we have seen that metrics based on temperature (Section 3.1.2) require a “packing” of the matrix that re-arranges the adjacency matrix’s rows and columns in such a way to maximize nestedness (i.e., minimize the nestedness temperature). Such metrics shift the question from whether a given matrix is significantly nested when its rows and columns are ordered by degree to whether a given matrix is significantly nested when its rows and columns are ordered in the way that maximizes nestedness. Answering the latter question requires accurate methodologies to produce the optimal ordering of rows and columns; presenting them is the main goal of this Section.

Historically, the first algorithm introduced with the goal of packing the matrix to maximize nestedness is the Nestedness Temperature Calculator (NTC) by Atmar and Patterson [22]. Despite some criticism [70,191], the algorithm has been widely used in the ecological literature in order to quantify the nestedness of spatial and interaction networks [48]. On the other hand, it suffers from shortcomings as its determination of the line of perfect nestedness (see Section 3.1.2) is not unique, and its iterative packing algorithm fails to identify the optimal packing in many empirical networks [70]. The genetic algorithm by Rodríguez-Girónés and Santamaría [70], called BINMATNEST, overcomes these limitations (Section 3.3.1) and it might be considered as the state-of-the-art approach to maximize nestedness in the ecological literature.

From a different angle, scholars have been interested in quantifying the competitiveness of nations and the sophistication of products in World Trade through network-based metrics [24,110]. These metrics take as input the binary adjacency matrix of the network that connects countries with the products they export. While this ranking problem is seemingly unrelated to the nestedness maximization problem in ecological networks, it turns out that one of the most studied country–product ranking algorithm, the fitness–complexity algorithm [24], produces rankings of countries and products that reveal a “triangular shape” [24] of the adjacency matrix of the country–product network. In other words, the fitness–complexity algorithm enhances the nestedness of the matrix. Section 3.3.2 reviews the algorithm and its variants [114–116]. A natural question emerges: is nestedness temperature better minimized by the genetic algorithm or by the fitness–complexity algorithm? Section 3.3.3 provides a comparison of the two approaches.

3.3.1. Genetic algorithms

The problem of sorting the rows and columns of the adjacency matrix in such a way to maximize nestedness is combinatorially difficult: there are indeed $N! M!$ possible permutation of rows and columns, which makes it essentially unfeasible to explore all the possible rearrangements and select the one with the lowest temperature. In similar situations, scholars have often resorted to genetic algorithms inspired by evolution [192]. In this class of computational models, one

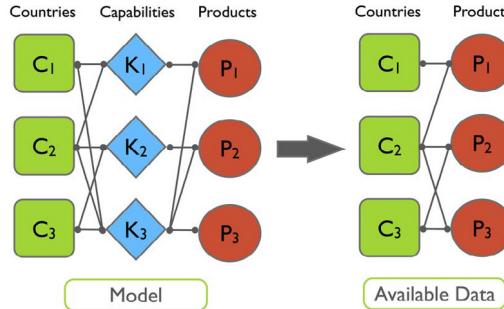


Fig. 9. The economic complexity interpretation of the bipartite country–product export network. In order to be produced and exported, each product requires a given set of capabilities. Only the countries that possess all the capabilities required to produce a given product can export it. The resulting observed bipartite network (right panel) can be therefore interpreted as a projection of a tripartite network where the capabilities layer cannot be observed. A model based on this simple scheme fits the structural properties of the country–product network (see Section 4.3.4).
Source: Reprinted from [193].

starts with a set of candidate solutions ("chromosomes") of the problem of interest, and recombines them according to pre-defined rules. Candidate solutions that provide a better solution to the problem – often, they exhibit larger values for a target function of interest – are given larger chance to "reproduce". The BINMATNEST algorithm by Rodríguez-Gironés and Santamaría [70] is precisely a genetic algorithm that rearranges the rows and columns of the adjacency matrix with the goal to minimize the nestedness temperature.

The BINMATNEST genetic algorithm starts from a set of candidate solutions $\{\mathbf{r}^R, \mathbf{r}^C\}$ which includes the ordering by degree, and several other orderings based on routines similar to the NTC algorithm. The algorithm proceeds by producing an "offspring" from a well-performing solution \mathbf{w} . To this end, the algorithm randomly selects a "partner" solution \mathbf{p} from the population of candidate solutions. Each element o_i of the offspring solution \mathbf{o} is given by w_i with probability 0.5, and it is otherwise determined by a combination of information from both \mathbf{w} and \mathbf{p} . This choice is performed independently for each solution, with the constraint that at least one offspring solution must be based on combination. The combination mechanism assigns $o_i = w_i$ for $i = 1, \dots, k$, where k is a random number uniformly extracted from $\{1, 2, \dots, N - 1\}$. For $i \in \{k + 1, \dots, N\}$, $o_i = p_i$ if that position had not been already assigned – i.e., if $p_i \notin \{w_1, \dots, w_k\}$; if $p_i \in \{w_1, \dots, w_k\}$, the element o_i is randomly extracted from all the positions that have not yet been assigned. Finally, each offspring solution suffers from a mutation with probability 0.1: a portion of the vector \mathbf{o} is randomly selected and its elements undergo a cyclic permutation.

The BINMATNEST algorithm produces matrices that exhibit significantly lower temperature than those produced by the NTC (see Figs. 4–5 in [70]). This has been verified in both synthetic and real networks; in real networks, the gap between the temperature by the NTC and the temperature by BINMATNEST tends to be larger for pollination networks than for spatial networks. Besides, while the temperature by BINMATNEST depends on network density (in particular, the lowest temperature values are achieved by very sparse or very dense networks), the probability that the observed temperature is attained by chance does not, which makes it legitimate to use the p -values of nestedness temperature to compare the degree of nestedness of systems of different size and density [70].

3.3.2. Non-linear iterative algorithms

Non-linear iterative algorithms (like the *fitness-complexity algorithm*) were not explicitly introduced with the goal to enhance the nestedness of a given matrix. The algorithms were originally aimed to rank countries and products in bipartite country–product networks where the countries are connected with the products they export [24]. Yet, they turn out to be effective in enhancing nestedness [71,73], which justifies their place here as methods to pack the adjacency matrix of a given network.

Fitness-complexity metric and its interpretation. The algorithm is motivated by simple economic considerations. The economic complexity¹² approach to macroeconomics views the export of a product as the result of a production process that requires from a country all the necessary capabilities in order to fabricate that product (see Fig. 9 for an illustration). In this view, countries that possess more capabilities are more competitive in the world trade, as they have the possibility to produce and export more products than countries with few capabilities.

Based on these lines of reasoning, Tacchella et al. [24] defined the *fitness*, F_i , of a given country i as the sum of the scores of the products exported by that country:

$$F_i = \sum_{\alpha} A_{i\alpha} Q_{\alpha}. \quad (37)$$

¹² We refer the interested reader to [https://www.mdpi.com/journal/entropy/special_issues/Economic_Complexity] for a recent special issue of the journal *Entropy* on economic complexity.

Table 5

Applications of the fitness–complexity algorithm and its variants to diverse systems. For each system, we provide a brief interpretation of both Fitness and Complexity score, and we refer to the mentioned references for all the details.

Type of network	System	Fitness score	Complexity score	Refs.
Economic	Country-product export	Country competitiveness	Product specialization	[24]
Economic	Country-food production	Country competitiveness	Food specialization	[195]
Knowledge production	Country-research field	Country competitiveness	Field complexity	[196]
Social	User-page engagement in Facebook	User engagement	Page impact	[197]
Ecological	Plant-pollinator	Pollinator importance	Plant vulnerability	[71]

Naturally, the products' scores $\{Q_\alpha\}$ depend on the exporting countries' scores. However, this dependence cannot be linear because a product that is exported by many countries is likely to require few capabilities to be produced, which means that it cannot be a sophisticated product.

Tacchella et al. [24] therefore observe that a product that is exported by specialist countries should be penalized. One of the simplest mathematical ways to enforce this idea is to define a given product's score as the reciprocal of the sum of the reciprocals of the fitness scores of its exporting countries:

$$Q_\alpha = \frac{1}{\sum_i A_{i\alpha} / F_i}. \quad (38)$$

According to this definition, if a product α is exported by a low-fitness country i , this country's small score F_i gives a large contribution $1/F_i$ to the sum in the denominator, which results in a small product score Q_α .

Eqs. (37)–(38) are not yet the final equations of the algorithm, because the countries' (products') scores depend on the products' (countries') scores, and we do not know either of them, a priori. In line with the widely-used power method to compute Google's PageRank [194], one sets a uniform initial condition [24]

$$\begin{aligned} F_i^{(0)} &= 1, \\ Q_\alpha^{(0)} &= 1, \end{aligned} \quad (39)$$

and then seeks to solve iteratively Eqs. (37)–(38):

$$\begin{aligned} \tilde{F}_i^{(n)} &= \sum_\alpha A_{i\alpha} Q_\alpha^{(n-1)} \\ \tilde{Q}_\alpha^{(n)} &= \frac{1}{\sum_i A_{i\alpha} / F_i^{(n-1)}}. \end{aligned} \quad (40)$$

At each step, the scores are further normalized by their mean:

$$\begin{aligned} F_i^{(n)} &= \tilde{F}_i^{(n)} / \langle \tilde{F}_i^{(n)} \rangle \\ Q_\alpha^{(n)} &= \tilde{Q}_\alpha^{(n)} / \langle \tilde{Q}_\alpha^{(n)} \rangle. \end{aligned} \quad (41)$$

Ideally, one would like to define the vector of country and product scores as the stationary point of these iterative equations. This is not always possible: one can show both numerically and analytically that for some shapes of a nested adjacency matrix, the scores of multiple countries and products converge to zero [114,115]. Conditional on the level of nestedness and the density of the adjacency matrix, such convergence can be as slow as a power-law of the number of iterations [114,115]. Of course, if many countries and products have zero scores, the resulting ranking cannot discriminate their relative importance. To bypass this issue, scholars have proposed various solutions. The simplest one is to halt the algorithm after a finite number of iterations, and to check a posteriori that the country and product scores are all larger than zero [72]. Other scholars [73,114] suggested to use a convergence criterion that relies on the convergence of the ranking of the nodes, and not on the convergence of their score.

Before introducing variants of the fitness–complexity algorithm, we stress the reason why the algorithm produces highly-nested adjacency matrices. The algorithm indeed not only rewards generalist countries and specialist products, but also ranks the products in such a way that a product's score is mostly determined by the score of the least-fit exporting countries. Therefore, given a pair of countries, the fittest one (i.e., in the ordered matrix, the one whose corresponding adjacency matrix's row lies above) is typically able to export additional products that are not exported by the least-fit country; these additional products tend to have a higher complexity score (i.e., their corresponding adjacency matrix's columns tend to lie more on the right) than those exported by the least-fit country, which is in agreement with the definition of nestedness.

We conclude this introduction to the fitness–complexity algorithm by emphasizing that the algorithm and its variants can be applied to any bipartite network (see Table 5). For example, Domínguez-García and Muñoz [71] have applied the

algorithm¹³ to mutualistic networks, with the goal to study its ability to identify structurally important nodes – more details are provided in Section 6.1.2. In mutualistic networks, countries and products are naturally replaced by active and passive species, respectively; the fitness and complexity scores represent the active species' importance and the passive species' vulnerability, respectively [71].

Generalized fitness-complexity algorithm. According to the fitness-complexity algorithm, the score of a product is largely determined by the scores of the least-fit countries. This dependence can be sharpened by introducing an exponent in the complexity definition [72,114]. By keeping the initial condition of the original algorithm (Eq. (39)) and the score normalization after each iteration (Eq. (40)), the equations of the resulting *generalized fitness-complexity algorithm* are [72,114]

$$\begin{aligned} F_i^{(n)}(\gamma) &= \sum_{\alpha} A_{i\alpha} Q_{\alpha}^{(n-1)} \\ Q_{\alpha}^{(n)}(\gamma) &= \frac{1}{\sum_i A_{i\alpha} / (F_i^{(n-1)})^{\gamma}}, \end{aligned} \quad (42)$$

where γ is a parameter of the method. The original fitness-complexity algorithm is obtained for $\gamma = 1$. Given the interpretation of the algorithm in terms of economic capabilities, we are mostly interested in the $\gamma > 0$ range. Increasing γ increases the dependence of product score on the fitness of the least-fit exporting country. This results in more nested adjacency matrices (see Fig. 3 in [72]), which proves to be beneficial for the identification of structurally important nodes, as we shall analyze in Section 6.1.2. At the same time, the rankings obtained with $\gamma > 1$ tend to be more sensitive to structural perturbations of the network's structure [72], which is a drawback especially for systems where a non-negligible fraction of the links might be unreliable, such as the World Trade [45,198]. A different generalization of the fitness-algorithm was introduced by Zaccaria et al. [197] to rank users' engagement and pages' impact in Facebook; they found that the resulting algorithm (which they called *PopRank*) can reliably predict the future activity of a Facebook page.

Minimal extremal metric. While the generalized fitness-complexity algorithm introduced in the previous paragraph allows us to fine-tune the dependence of a product's score on the score of the least-fit exporting countries, it is instructive to consider the limit case where the products' score is *entirely* dependent of the least-fit exporting country's score. By keeping the initial condition of the original algorithm (Eq. (39)) and the score normalization after each iteration (Eq. (40)), the equations of the resulting *minimal extremal metric* are [115]

$$\begin{aligned} F_i^{(n)} &= \sum_{\alpha} A_{i\alpha} Q_{\alpha}^{(n-1)} \\ Q_{\alpha}^{(n)} &= \min_{i: A_{i\alpha}=1} \{F_i^{(n)}\}. \end{aligned} \quad (43)$$

This metric corresponds to the limit $\gamma \rightarrow \infty$ of the generalized fitness-complexity metric, and it is interesting for two main reasons: (1) for empirical networks, with respect to the FCM, it improves the nested packing of the adjacency matrix (see Fig. 4 in [115]); (2) for perfectly nested networks, the vector of node scores is related to the network's degree sequence through a simple, exact mathematical relation (see Eq. 11 in [115]), which allows us to build a simple intuition on the convergence properties of non-linear ranking algorithms.

A variant with improved convergence properties. To improve the convergence properties of the fitness-complexity algorithm, Stojkoski, Utkovski and Kocarev [116] defined a modified fitness-complexity algorithm. By keeping the initial condition of the original algorithm (Eq. (39)) and the score normalization after each iteration (Eq. (39)), the equations of their proposed algorithm (hereafter referred to as SUK fitness-complexity algorithm, after its authors) read [116]

$$\begin{aligned} F_i^{(n)} &= \sum_{\alpha} A_{i\alpha} Q_{\alpha}^{(n-1)} \\ Q_{\alpha}^{(n)} &= \frac{1}{\sum_i A_{i\alpha} (N - F_i^{(n-1)})}. \end{aligned} \quad (44)$$

The r.h.s. of the second line of Eq. (44) can be viewed, approximately, as a second-order expansion of the original fitness-complexity algorithm.¹⁴ Importantly, product and country scores cannot converge to zero in the SUK algorithm. This is a simple consequence of the fact that

$$\sum_i A_{i\alpha} (N - F_i^{(n-1)}) = N k_{\alpha} - \sum_i A_{i\alpha} F_i^{(n-1)}, \quad (45)$$

¹³ To stress the different interpretation of the scores by the algorithm, the fitness-complexity algorithm is dubbed as “MusRank” by Domínguez-García and Muñoz [71].

¹⁴ This can be seen by expanding the inverse $1/F$ of country score around $F^* = N/F$, where f_0 is a real number. We obtain that $1/F \approx 2f(N - fF/2)/N^2$. Therefore, each contribution to product score can be expressed, approximately, in terms of the difference between the total number of countries N and a term proportional to country fitness F . In the special case $f = 2$ ($F^* = N/2$) [116], each contribution becomes precisely equal to $N - F$, i.e., the term used in the second line of Eq. (44).

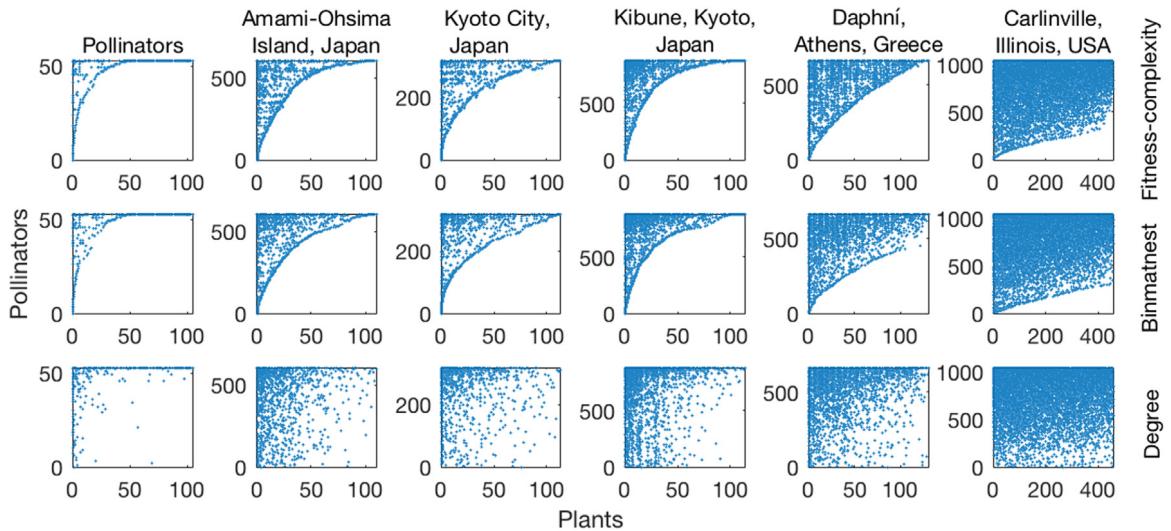


Fig. 10. Interaction matrices of six mutualistic bipartite networks when rows and columns are ranked by the fitness-complexity algorithm (upper panels), by the BINMATNEST algorithm (intermediate panels), and by degree (lower panels). The matrices ranked by the fitness-complexity and BINMATNEST algorithms exhibit a higher degree of nestedness than those packed by degree.

Source: Reprinted from [73].

which implies

$$Q_\alpha^{(n)} = \frac{1}{\sum_i A_{i\alpha} (N - F_i^{(n-1)})} \geq \frac{1}{N k_\alpha} > 0. \quad (46)$$

The algorithm was used by Stojkoski et al. [116] to assess the impact of services on the rankings by economic complexity metrics – a similar goal was also pursued through the fitness-complexity algorithm by a recent paper [199] co-authored by World Bank members. Stojkoski et al. [116] found that services tend to be ranked higher than goods in product rankings, and high-fitness countries tend to have more developed service sectors. By using the fitness-complexity algorithm, Zaccaria et al. [199] found that including services trade in the economic fitness approach can substantially alter the rankings of countries and products.

3.3.3. Nestedness temperature minimization: BINMATNEST or fitness-complexity?

So far, we have independently introduced two approaches – genetic algorithms and iterative non-linear ranking algorithms – to rank the nodes of a given matrix. We now ask ourselves: how do the two methods perform in minimizing nestedness temperature or, equivalently, in maximizing the adjacency matrix's degree of nestedness? Preliminary results by Domínguez-García and Muñoz [71] suggested that the fitness-complexity algorithm "packs the matrices" substantially better than the widely-used NTC (see Fig. 5 in [71]). They concluded that the fitness-complexity algorithm "should be used (rather than existing ones) to measure nestedness in bipartite matrices" [71]. On the other hand, shortcomings of the NTC were already pointed out by Rodríguez-Gironés and Santamaría [70], which motivated them to introduce the genetic algorithm described in Section 3.3.1.

How does the fitness-complexity algorithm perform when compared with the BINMATNEST algorithm? A detailed comparison of the matrices packed by BINMATNEST with those packed by the fitness-complexity algorithm has been performed recently [73]. Lin et al. [73] found that the matrices as ranked by fitness-complexity and BINMATNEST are substantially better "packed" than those ranked by degree (see Fig. 10). More surprisingly, they found that the temperature of the matrices ranked by the fitness-complexity algorithm is lower than that of the matrices ranked by BINMATNEST for the majority of the mutualistic bipartite networks that they analyzed. The only networks where BINMATNEST turned out to substantially outperform the fitness-complexity algorithm were characterized by small size and high density.

Lin et al. [73] concluded that beyond its application in trade networks, the fitness-complexity algorithm has the potential to become a standard tool in nestedness analysis. It remains open to assess whether variants of the fitness-complexity algorithm can further reduce nestedness temperature, and the impact of improved algorithms for nestedness minimization on the acceptance or rejection, based on suitable null models, of the hypothesis that a network is nested.

3.4. Bottom-line: How to measure nestedness?

Scholars have introduced several distinct metrics to measure nestedness. Some of them (like NODF [150] and spectral radius [83]) measure the level of nestedness by degree, whereas others (like the nestedness temperature by the NTC [22],

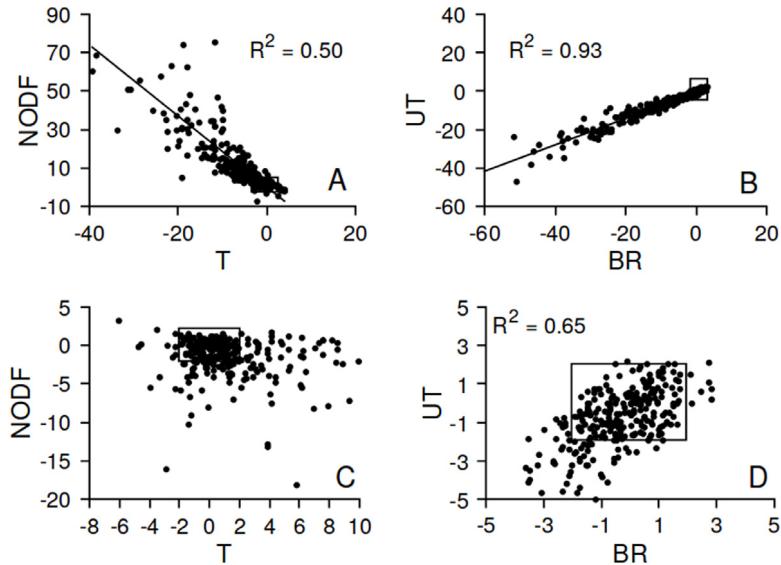


Fig. 11. Pairwise comparisons of the z-scores by four different nestedness metrics (NODF, nestedness temperature T , number of discrepancies BR , number of unexpected transformations UT – see Section 3 for all the definitions) for the spatial datasets collected by Atmar and Patterson (see Appendix A). The z-scores are based on the EE model which only preserves network size and density (panels A–B), and the FF model which preserves exactly the degree sequence (panels C–D). The squares delimit the regions where nestedness is not significative according to any of the two metrics. Source: Reprinted from [48].

by BINMATNEST [70], and by the fitness-complexity algorithm [73]) measure the maximal level of nestedness in the system, and they require a reordering of the rows and columns of the (bi)adjacency matrix.

Generally, these metrics depend on basic systemic properties such as network size, density, and degree distribution. This makes it necessary to assess the statistical significance of their observed values based on a null model. How to choose the null model is controversial: scholars have warned against the risk of both an excessive number of false positives, if the null model is too loose (e.g., Equiprobable–Equiprobable model), or an excessive number of false negatives, if the null model is too conservative (e.g., Fixed–Fixed model).

There is no universal answer on which combination of nestedness metric and null model should be adopted to investigate the presence and implications of nestedness [9]. A valid operational strategy is to always make sure that obtained results on nestedness hold for different metrics and similar null models. Information about publicly available software for the implementation of nestedness metrics together with null models for statistical significance tests is provided in Appendix A.

4. Relation between nestedness and other systemic properties

Nestedness is a network structural property, which means that it depends on the adjacency matrix A of the network of interest. Two questions emerge: how is this property related to other known network properties (such as degree distribution, assortativity, modularity, etc.)? How is the property related to important properties of real systems that are not directly related to network structure? Section 4.1 addresses the first question, revealing a tight bound between nestedness and the network's degree distribution [169] and disassortativity [146]. Besides, nestedness turns out to be a generalization of the core–periphery structure [41], and it has a multifaceted relation with modularity [32,200].

Section 4.2 discusses the relation between nestedness and ecological properties that are not captured by the topology of interactions. These properties include macroecological properties, species relative abundances, and forbidden links [9,68]. Section 4.3 focuses on economic systems, and it describes the relation between nestedness and various economic properties. We mostly focus on the predictive power of nestedness for the success or failure of firms [28], the appearance and disappearance of links in spatial and trade networks [29], the future economic development of countries [122], and on the relation between nestedness and hidden capabilities in world trade [29,52].

4.1. Nestedness and other network properties

This Section studies the relation between nestedness and other network properties: degree distribution (Section 4.1.1), disassortativity (Section 4.1.2), core–periphery structure (Section 4.1.3), modularity (Section 4.1.4). For each of these network properties, we will describe its relevance in network analysis, introduce the main metrics to measure it, and discuss its relation with nestedness.

4.1.1. Nestedness and degree distribution

A natural question which invariably arises for any network structural pattern is whether the property can be simply explained by the network's degree distribution [11] or by other local, higher-order network properties [201]. Almost as soon as the first instances of nested structures were reported [21], ecologists have started investigating whether the observed nested patterns could be explained by the degree distribution alone. The question has been addressed by means of both the degree-preserving randomization procedures introduced in Section 3.2.1 (e.g., PP and FF model), and by the maximum-entropy techniques described in Section 3.2.2.

Biogeographic networks. As nestedness was first introduced in the biogeography literature [21], the first attempts to study the relation between network nestedness and degree sequence were carried out for bipartite species-island networks. For spatial networks, Brualdi and Sanderson [154] found that the observed number of discrepancies (a metric for nestedness introduced in Section 3.1.1) in 33 empirical networks is compatible with the number of discrepancies observed in randomized networks generated with the FF model. They argued that the main reasons why nested patterns were widely found in previous studies [74,202] was that “violation of the row or column sums (or both) did not sufficiently constrain the sample space containing the nested species subsets”.

Ulrich et al. [48] performed an extensive analysis on the 286 empirical networks previously collected by Patterson and Atmar (see Appendix A). They focused on the FF model which preserves the degree sequence of both kinds of nodes in bipartite networks (see Section 3.2.1). They both measured the correlation between the z-scores of different nestedness metrics (NODF, number of discrepancies, temperature, unexpected number of transformations U_T) and assessed the number of empirical networks that exhibit a statistically significant nestedness (see Fig. 11). Out of the 286 empirical networks, they found 113 of them to be significantly nested according to at least one of the considered nestedness metrics. Intriguingly, they found only 11 networks to be significantly nested according to all the four considered nestedness metrics. Ulrich et al. [48] also found that the EE null model (i.e., a null model that only preserves network size and density, see Section 3.2.1) produces high values of z-score for most of the networks, regardless of the adopted nestedness metric: the choice of the null model has a strong impact on our conclusions on the significance of a nested pattern. As expected, more constrained models lead to fewer significantly-nested networks.

Ecological interaction networks. Already the first work by Bascompte et al. [23] found that for 20% of the analyzed pollination networks, nestedness can be explained by a PP null model that preserves, on average, animals' and plants' degree; they also found that this percentage drops to zero when restricting the analysis to sufficiently large networks. Joppa et al. [67] also found that most mutualistic networks are more nested than one would expect by degree sequence alone. They pointed out that the networks that are more nested than expected by degree sequence are typically the large ones. From a different perspective, Medan et al. [80] computed analytically the degree sequence associated with a particular line of perfect nestedness.

More recently, Borràs et al. [169] used the maximum-entropy approach for bipartite networks (see Section 3.2.2) to assess the relation between nestedness and degree sequence in 167 mutualistic networks. Their conclusions are different from those by previous studies [23,67]. Based on NODF and spectral radius, they found that among all the analyzed networks, only a tiny fraction of them exhibit z-score larger than two. They concluded that in mutualistic networks, the “observed nested structure of the ecological communities studied is, in fact, a mere consequence of the degree sequences of the two guilds”. They interpreted this result as a consequence of the fact that for networks with highly-heterogeneous degree distribution, disassortative structures tend to have larger entropy than non-disassortative ones [203]. This happens because a specialist has many more possibilities to be connected with a generalist than with another specialist. As nestedness and disassortativity are strongly related (see Section 4.1.2), it follows that for a highly-heterogeneous degree distribution, maximizing entropy can lead to nested structures [169].

World Trade networks. Results on the significance of nestedness in World Trade are intriguing. Using the maximum-entropy approach described in Section 3.2.2, Saracco et al. [112] quantified the significance of the nestedness (as measured by NODF, see Section 3.1.3) of the bipartite country–product export network (NBER dataset) over the period 1963–2000. The authors found that the z-scores of the network's NODF (determined numerically through the maximum-entropy framework described in Section 3.2.2) are always smaller than two. They concluded that the degree sequence allows us to reproduce the nestedness of the World Trade bipartite network. An analysis of the World Trade over the period 1995–2010 (see Fig. 12) shows again that network NODF is explained by the network degree sequence [26]. Nevertheless, the row-contribution to NODF (see Section 3.1.3) exhibits statistically significant values (z-score above two) from 1995 to 2002. The z-score declines from 1995 to 2002, lying on an almost constant plateau below $z = 2$ over the period 2003–2010. Based on similar findings for other structural properties, Saracco et al. [26] argued that the increasing randomness of the world trade network over the 2003–2007 period might be interpreted as an early sign of the 2007–2008 financial crisis.

Nestedness: Significant pattern or a consequence of the degree sequence? The question is tightly related to the choice of the null model and its implications. Should one prefer unconstrained models (such as the EE model, see 3.2.1) where the degree of the nodes is allowed to vary, or should one prefer a constrained null model (like the FF model, see 3.2.1) where the nodes' degree is fixed exactly (like in the FF model, see 3.2.1) or on average (like in the PP model and in the

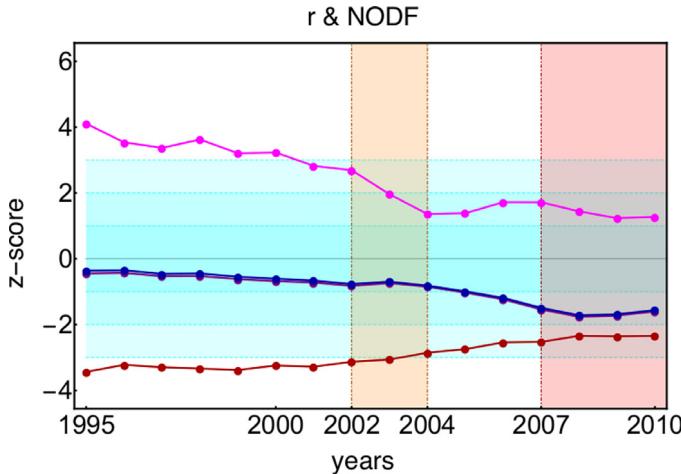


Fig. 12. Temporal evolution in the World Trade for the z-score of three quantities: NODF (blue dots), row-contribution to NODF (pink dots), assortativity (brown dots). The NODF and its contributions are based on the bipartite country–product export network; the z-scores are based on the maximum entropy model described in Section 3.2.2 which preserves, on average, the nodes’ degree. The row-contribution to NODF (i.e., the country-contribution) shows a non-trivial trend: it is significant in the first years but it gradually declines, falling below the $z = 2$ line after 2003. Assortativity shows a monotonous trend as well, yet it remains significantly negative ($z < 2$) over the whole observation time span.

Source: Adapted from [26].

maximum-entropy approach, see Sections 3.2.1–3.2.2)? Ecologists have widely debated this question. We can identify two main viewpoints in the literature.

Scholars have warned against the possible overestimation of Type-I and Type-II errors for unconstrained and constrained models, respectively. It has been widely recognized [48] that unconstrained models where all species have the same interaction probability might make it too easy for a network to achieve a statistically significant degree of nestedness. In other words, given the null hypothesis that the null model explains the observed degree of nestedness, one risks rejecting the null hypothesis when it is true (Type-I error [62]). A long stream of works [59,112,154,169] have employed constrained null models where the degree of the nodes is fixed, exactly or on average. As a result, the fact that a null model that preserves the degree sequence can generate networks of statistically comparable nestedness as the real ones is known since long time [154].

On the other hand, it has been also recognized that in constrained models (such as the FF model), fixing exactly the degree sequence might make it too hard for a network to achieve a statistically significant degree of nestedness. In other words, given the null hypothesis, one risks accepting the null hypothesis when it is false (Type-II error [62]). Some scholars [204] advocated the use of models that do not constrain the degree sequence of both kinds of nodes, based on the following argument, summarized by Gotelli and Ulrich [62]: with constrained models, “if the biological processes (e.g. competition) affect the constrained elements (e.g. matrix row totals) then the effect of interest has been smuggled into the test, which reduces the sample space and leads to excessive type II errors”. Because of the small size of the sample space, constrained models can lead to paradoxes: for example, a perfectly nested matrix might be classified as maximally non-nested by the FF model if it is the only possible matrix with that degree sequence (see the Supplementary Information in [83]).

This debate points out the importance of considering carefully the properties of the ensemble of random graphs used to assess the significance of observed patterns: classifying a network as non-significantly nested might be a consequence of the limited number of networks in the ensemble of networks with the same (or similar) degree sequence. A detailed analysis of Type I and Type II errors for both constrained and unconstrained null models can be found in [61]: they found that in a set of non-random matrices that contains nestedness, the FF model rejected the null hypothesis only less than 10% of the times for most of the nestedness metrics. In light of this discussion and recent developments [64], we recommend the detailed analysis of Type I vs Type II errors as a necessary step to be performed before adopting a given null model for the analysis. If a null model is not able to classify perfectly or almost perfectly nested networks as significantly nested, concluding that nestedness is not significant based on its results might be misleading. A similar analysis might also explain the reasons behind the contrasting results in previous literature [23,48,67,169].

4.1.2. Nestedness and disassortativity

A network exhibits nestedness if the neighborhood of a node is contained in the neighborhoods of the nodes with higher degrees. A direct consequence is that low-degree nodes tend to only interact with high-degree nodes, whereas high-degree nodes interact with both other high-degree nodes and with low-degree nodes. As a result, the average degree of the nodes that interact with high-degree nodes tends to be lower than that of the nodes that interact with low-degree nodes.

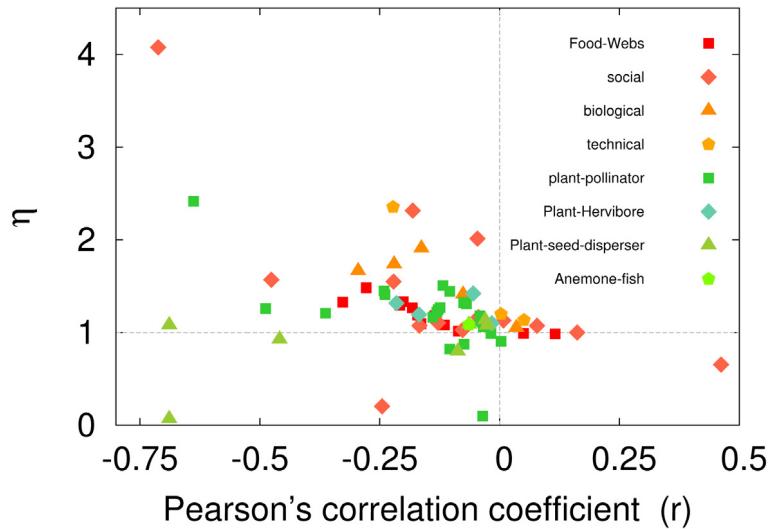


Fig. 13. The relation between nestedness (as measured by the JDM-NODF η defined by Eq. (19)) and disassortativity (as measured by the linear correlation r between the degrees of the two nodes that belong to each link) for empirical networks of various kinds. The most nested networks (large η) are all disassortative ($r < 0$), whereas the nestedness of assortative networks ($r > 0$) tends to be comparable with that of random networks ($\eta \approx 1$).

Source: Reprinted from [146].

This property is typically referred to as *disassortativity* in network science [11,205]. Therefore, nestedness is expected to be significantly correlated with plausible metrics for disassortativity; this expectation has been confirmed by analysis of synthetic and empirical networks [146,206]. Below, we provide more details on these results.

Assortativity and disassortativity. Assortativity in networks generally refers to the tendency of nodes to connect with nodes that are similar to them. By contrast, *disassortativity* refers to the tendency of nodes to connect with nodes that are different from them (see [11], Section 7.13). Of course, such general definitions strongly depends on which properties we consider when measuring the similarity between two given nodes. There can be two types of assortativity: assortativity by enumerative node properties (e.g., measuring the tendency of students in a given school to establish friendships with students of the same gender) and assortativity by scalar node properties (e.g., measuring the tendency of students in a given school to establish friendships with students of similar age or parental income) [11].

From a network perspective, it is instructive to measure the assortativity by degree, which aims to quantify the tendency of nodes to connect with nodes of similar degree [10,11]. Assortativity by degree has been found in systems as diverse as social networks [207], scientific co-authorship networks [205], paper [208] and patent [209] citation networks, World Wide Web [210], among many others. On the other hand, disassortativity has been found in metabolic networks [10], the Internet [211], food webs [205,210], among many others.

To measure the degree of assortativity in a given network, a commonly employed metric is the Pearson's linear correlation coefficient r between the degrees of pairs of nodes that belong to a given edge. Assortative (disassortative) networks are characterized by positive (negative) values of r , meaning that the degree of a link's node is positively (negatively) correlated with the degree of the other node that forms the link.

Nestedness and disassortativity. Abramson et al. [206] found that perfectly nested network exhibit negative assortativity coefficient r . The coefficient decreases (i.e., the network becomes more disassortative) as network density increases. Johnson et al. [146] showed that network disassortativity is significantly correlated with nestedness, as their metric η for nestedness (defined by Eq. (24)) is negatively correlated with the Pearson's linear correlation coefficient r between the degrees of pairs of nodes that form the network links. They showed this both in synthetic networks with tunable assortativity coefficient (see Fig. 2 in [146]) and in 60 unipartite and bipartite empirical networks of diverse nature (see Fig. 13). They concluded that “disassortative networks are typically nested and nested networks are typically disassortative”.

4.1.3. Nestedness and core-periphery structure

A core-periphery structure is a network structure composed of a “core” of nodes that are connected with all the other nodes, and a “periphery” of nodes that tend to be only connected with the nodes in the core. Such structure has been investigated and observed in a wide variety of systems, including international trade [212,213], social networks [214,215], human brain network [216,217], mutualistic networks [218,219], R&D networks [220], financial networks [129], among

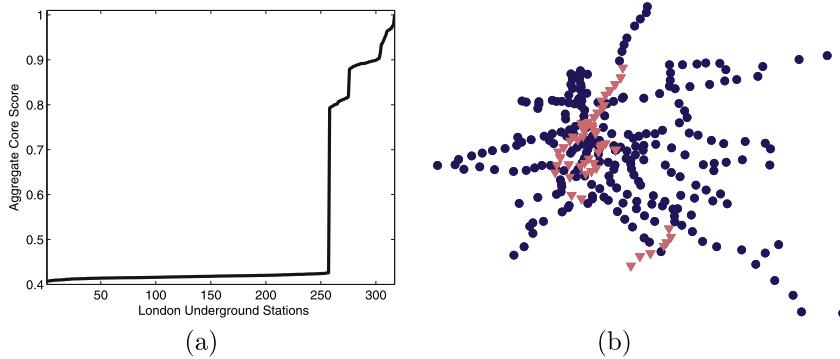


Fig. 14. Core nodes and core–periphery structure of the London Underground stations. The node-level aggregate core score C (defined in Eq. (50)) reveals that there are 60 important stations in the system (panel (a)). The structural importance of these stations (pink triangles) compared to the non-core stations (blue circles) can be appreciated in panel (b), where the stations' position is determined by their geographical location. The two clusters of cores correspond to central stations in London (north cluster) and to the stations close to Waterloo station (south cluster).
Source: Reprinted from [224].

many others. We refer the interested reader to [221] for a recent review on the structure and dynamics of core–periphery networks. Some scholars have observed that compared with the community detection problem (see [222,223] and Section 4.1.4), “other types of mesoscale structures [...] have received much less attention than they deserve” [224], and that the core–periphery structure is one of those.

In a core–periphery structure, the peripheral nodes’ neighborhoods are included by construction in the core nodes’ neighborhoods; therefore, it is natural to conjecture a relation between such structure and nestedness. This relation has been explicitly found by Lee [41] in both empirical and synthetic networks. It suggests that the methods and implications of nestedness analysis might be highly relevant to all systems where core–periphery structures are typically found.

Detection of core–periphery structures. Before investigating the relation between nested and core–periphery structures, we provide a brief introduction to the methods to detect core–periphery structures in networks. We describe the Borgatti–Everett (BE) approach [214] and its subsequent generalizations [41,224] that have been used to investigate the relation between core–periphery structures and nestedness [41]. We start by introducing the Borgatti–Everett (BE) approach [214] to detect a single core–periphery structure in unipartite networks.¹⁵ To identify core and peripheral nodes, Borgatti and Everett introduced, for each node i , a binary variable x_i such that $x_i = 1$ and $x_i = 0$ for core and peripheral nodes, respectively. They also considered an ideal core–periphery structure where the core nodes are connected with all the nodes, whereas the peripheral nodes are only connected with the core nodes. In terms of the \mathbf{x} variables, the corresponding adjacency matrix A^{CP} can be simply written as $A_{ij}^{CP} = x_i + x_j - x_i x_j$ – i.e., $A_{ij}^{CP} = 1$ if i or j belong to the core, whereas periphery–periphery links are strictly forbidden. A simple metric to quantify the similarity between the network’s adjacency matrix, A , and the ideal core–periphery structure, A^{CP} , is the quality function

$$Q^{CP} = \sum_{(i,j)} A_{ij} A_{ij}^{CP} (\mathbf{x}). \quad (47)$$

To distinguish among core and peripheral nodes, one seeks to find the vector \mathbf{x} that maximizes Q^{CP} .

Borgatti and Everett [214] recognized that the assumption of a binary distinction between core and peripheral nodes might be simplistic to accurately describe real systems’ nodes. To make the model more flexible, they considered a continuous variant where each node is endowed with a “coreness” value c_i . The elements of the ideal core–periphery network’s adjacency matrix are then defined as $A_{ij}^{CP} = c_i c_j$. Such matrix elements: (1) are large when both i and j have a high coreness value; (2) are small when both i and j have a small coreness value; (3) assume intermediate values when one of the two nodes has a small coreness, the other one has a large coreness value. The maximization of the resulting quality function [214]

$$Q^{CP} = \sum_{(i,j)} A_{ij} c_i c_j \quad (48)$$

gives as output the optimal vector of coreness scores. Interestingly, as recognized by Borgatti and Everett [214], this optimal vector is equivalent to the Bonacich eigenvector centrality [228].

¹⁵ While we will focus on the Borgatti–Everett method in this Section, we refer the interested reader to [44,224–227] for alternative methods for the detection of core–periphery structures.

Rombach et al. [224] generalize this approach by introducing a parameterization $c_i(a, b)$ of the core score in terms of two parameters a and b . They chose a “sharp” function that separates the core from the peripheral nodes:

$$c_i(a, b) = \begin{cases} \frac{i(1-a)}{2[bN]} & \text{if } i \in \{1, \dots, [bN]\}, \\ \frac{(1-[bN])(1-a)}{2(N-[bN])} + \frac{1+a}{2} & \text{if } i \in \{[bN] + 1, \dots, N\}. \end{cases} \quad (49)$$

The parameter $b \in [0, 1]$ determines the size of the core (all the nodes are in the core when $b = 0$, whereas the core is empty when $b = 1$). The parameter a determines the score gap between the lowest-score core node and the highest-score peripheral node. When $a = 1$, such score gap is maximal, c_i is a discontinuous function, and the nodes are either core nodes ($c_i = 1$) or peripheral nodes ($c_i = 0$), without intermediate coreness values. In line with Borgatti and Everett [214], the assignment of the nodes to the core or the periphery of the network is performed in such a way to maximize the core-periphery function $Q^{CP}(a, b) = \sum_{(i,j)} A_{ij} c_i(a, b) c_j(a, b)$. The aggregate core score of node i is defined as the weighted average of $c_i(a, b)$ over all the possible values for the parameter pair (a, b) ; parameter pairs (a, b) that lead to larger values of the core-periphery quality function Q^{CP} give larger contribution to $C_i(a, b)$:

$$C_i(a, b) = Z \sum_{(a,b)} c_i(a, b) Q^{CP}(a, b), \quad (50)$$

where the normalization factor Z ensures that $\max_i C_i = 1$. The aggregate core score of a node is informative about its centrality in the network (see Fig. 14).

Eq. (49) can be generalized to bipartite networks [41]. This is done by introducing two vectors $\{c_i^R(a^R, b^R)\}$ and $\{c_\alpha^C(a^C, b^C)\}$ of coreness scores for row-nodes and for column-nodes, respectively. The model has now four parameters in total: two parameters, $\{a^R, b^R\}$ and $\{a^C, b^C\}$, parametrize the coreness score of row-nodes and column-nodes, respectively. In line with [224], one can again define the aggregate core score of row-node i and column-node α as [41]

$$\begin{aligned} C_i^R &= Z^R \sum_{(\mathbf{a}, \mathbf{b})} c_i(a^R, b^R) Q^{CP}(\mathbf{a}, \mathbf{b}), \\ C_\alpha^C &= Z^C \sum_{(\mathbf{a}, \mathbf{b})} c_\alpha(a^C, b^C) Q^{CP}(\mathbf{a}, \mathbf{b}), \end{aligned} \quad (51)$$

where $(\mathbf{a}, \mathbf{b}) = (a^R, b^R, a^C, b^C)$, and we defined the quality function

$$Q^{CP}(a, b) = \sum_{i, \alpha} A_{i\alpha} c_i(a^R, b^R) c_\alpha(a^C, b^C). \quad (52)$$

The degree of core-periphery organization ξ of the network can be defined as [41]

$$\xi = \frac{\sum_{i, \alpha} A_{i\alpha} C_i^R C_\alpha^C}{E \sum_i C_i^R \sum_\alpha C_\alpha^C}. \quad (53)$$

The numerator of this expression has the same form as Eq. (50), with the difference that it uses the aggregate core scores which do not depend on the specific (\mathbf{a}, \mathbf{b}) parameters.

Nestedness and core-periphery structure. Lee [41] compared the NODF nestedness metric (defined by Eq. (14)) with the core-periphery metric ξ (defined by Eq. (53)) on 89 mutualistic bipartite networks (both plant–pollinator and seed dispersal) and in synthetic networks with tunable nestedness. In mutualistic networks, he found a strong correlation between NODF and ξ ; however, for those systems, both NODF and ξ turn out to be strongly correlated with both edge density (see Fig. 5 in [41]). To factor out the impact of edge density on the NODF- ξ correlation, Lee considered a model where nestedness can be tuned while preserving the network’s edge density. Such analysis reveals again a strong correlation between NODF and ξ , but both metrics turn out to be strongly correlated with the variance of the degree distribution (see Fig. 6 in [41]). To factor out the impact of both edge density and degree heterogeneity on the NODF- ξ correlation, Lee [41] considered a randomization procedure (analogous to the FF model, see Section 3.2.1) that preserves exactly the degree sequence. Again, NODF and ξ are strongly correlated, which suggests that while not equivalent, nestedness and core-periphery are two closely-related network properties, regardless of the network’s degree distribution.

4.1.4. Nestedness and modularity

Modularity aims to quantify how well a network can be partitioned into different groups of nodes (referred to as *modules*) such that nodes that belong to the same group are more likely to be connected than nodes that belong to different groups. In the network science language, the identification of modules can be seen as a technique of community detection¹⁶ [223].

¹⁶ A large number of community detection methods have been introduced in the literature, involving techniques as diverse as random-walk minimum description length [229], statistical inference techniques [230,231], spectral methods [232], among many others. To quote Hric and Fortunato, “as long as there will be networks, there will be people looking for communities in them” [223]. We refer to [222,223] for comprehensive reviews of community detection techniques.

It turns out that the relation between nestedness and modularity strongly depends on network connectance: low-density (high-density) networks exhibit a positive (negative) correlation between nestedness and modularity (see Fig. 15 below and the related discussion). Beyond this correlational evidence, it is interesting to study the joint dynamics of these two properties. Borge-Holthoefer et al. [32] recently found that a social system can transition from a modular to a nested structure as a consequence of a special event (see Fig. 16 below). Besides, a module can exhibit an internal nested structure, and one can design structural functions to detect subsets of nodes that exhibit a nested pattern of interaction; this topic will be addressed in Section 7.

Modularity. Modularity [233] is one of the most studied properties of complex networks. The modularity score Q of a given network quantifies how well the network can be partitioned into blocks (or communities, or compartments) that have a high internal density of links and few connections with other blocks. For a undirected unipartite network, to compute the modularity of a given partition $\Xi = \{\Xi_i\}$ of the network into blocks – Ξ_i denotes here the block to which node i is assigned to – we compare the number of edges A_{ij} between two nodes i and j that belong to the same community with the expected number of edges E_{ij} according to a suitable null model. In formulas, for an undirected network, the modularity Q is defined as

$$Q = \frac{1}{2E} \sum_{i,j} (A_{ij} - E_{ij}) \delta(\Xi_i, \Xi_j), \quad (54)$$

where $E_{ij} = k_i k_j / (2E)$ is the expected number of edges between i and j under the configuration model (or, equivalently, the PP null model), and the delta function $\delta(\Xi_i, \Xi_j)$ restricts the sum to the pairs of nodes that belong to the same community.

Modularity can be readily generalized to directed [234] and bipartite networks [235]. The most popular extension of modularity to bipartite networks is arguably Barber's modularity [235]. In Barber's modularity, one uses the fact that in bipartite networks, nodes of one type can only connect with nodes of the other type. Therefore, only pairs of nodes of dissimilar type contribute to the bipartite modularity Q :

$$Q = \frac{1}{E} \sum_{i,\alpha} (A_{i\alpha} - E_{i\alpha}) \delta(\Xi_i, \Xi_\alpha), \quad (55)$$

where $E_{i\alpha} = k_i k_\alpha / E$ is the expected number of edges between i and α according to the configuration model (PP model, in the language of Section 3.2.1). Heuristic algorithms developed to maximize the modularity function for unipartite networks can be extended to maximize Barber's modularity as well.

Modularity optimization is one of the most popular community detection techniques. Research on modularity has focused on evaluating its ability to reconstruct ground-truth communities in synthetic networks [236–238], designing fast and effective heuristics to optimize it [239,240], unveiling its incapability to detect small modules under particular circumstances ("resolution limit") [241], quantifying its statistical significance [242], generalizing it to multilayer and temporal networks [243], among other problems.

Nestedness and modularity in ecological networks. Two natural questions arise: Do ecological networks exhibit modularity? If yes, what is the relation between nestedness and modularity? Early studies attempting to uncover the modular ("compartmentalized") structure of ecological networks date back to the 80s [244], and related studies followed in the 90s and early 2000s [245,246]. These works adopted simple statistics to quantify the "compartmentalization" of the system, and concluded that empirical food webs [245] and mutualistic networks [246] can be compartmentalized. More recently, based on the modularity function defined above, Olesen et al. [200] evaluated the modularity (optimized through a simulated annealing algorithm) of 51 pollination networks. They found that all networks of more than 50 species were significantly modular, whereas all networks with less than 50 species were not significantly modular. In addition, they found no significant correlation between modularity and nestedness temperature. They argued that nestedness emerges from the assembly of distinct modules "glued together by interactions among modules".

Fortuna et al. [39] analyzed 95 ecological communities including plant-animal mutualistic networks and host-parasite networks. First, they found a significant correlation between nestedness and modularity for plant-pollinator networks ($r = 0.363, p = 0.035$) but not for plant-seed disperser ($r = 0.151, p = 0.503$) and host-parasite ($r = 0.066, p = 0.689$) networks. Importantly, by using two statistical null models, they found that the relation between nestedness and modularity strongly depends on network connectivity: at low (high) connectivity, the two properties tend to be positively (negatively) correlated – in other words, the correlation between the nestedness-modularity correlation and network connectance is negative (Fig. 15).

Beyond the overall correlation between nestedness and modularity, another important question is the quantification of the level of nestedness within the modules. While one may naively compute the level of nestedness within the modules detected by modularity-optimization algorithms, recent studies [40] suggest that this approach can lead to misleading results. This problem is the main topic of Section 7.

Besides structural analysis, it is natural to investigate how such a compartmentalized structure affects dynamical processes on the network. The question is typically addressed by means of dynamical models and stability analysis [247,248]. Recently, Gilarranz et al. [249] tackled this question with an experimental setup that allowed them to study the secondary

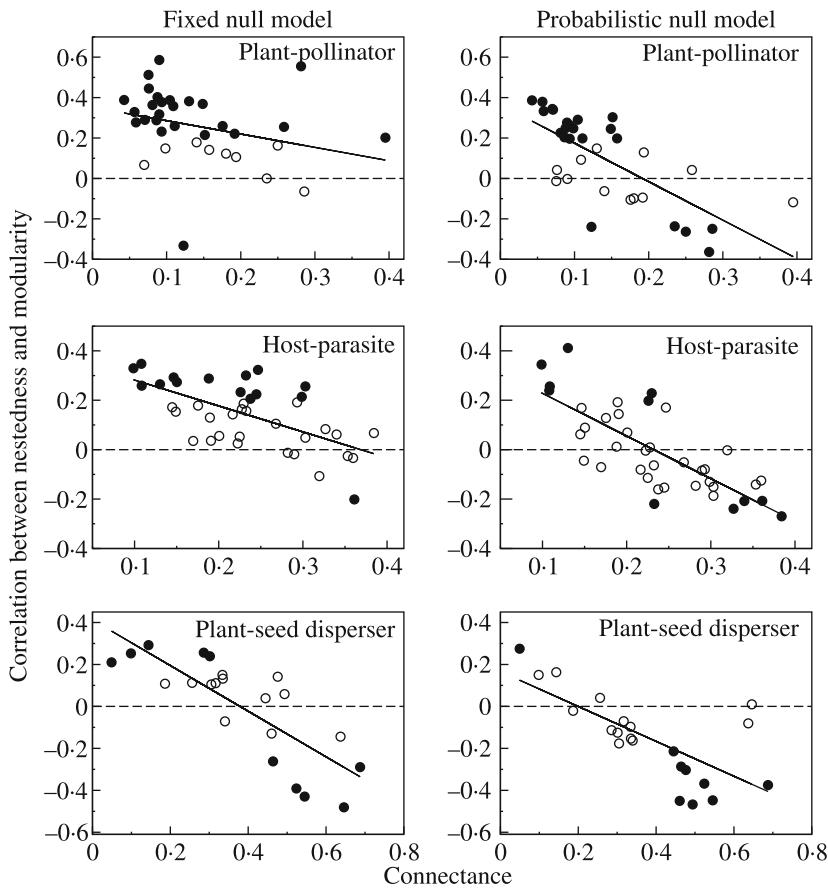


Fig. 15. Nestedness-modularity correlation as a function of network connectance – each dot represents a network. The nestedness and modularity values represented here are averages over 100 randomizations of the original networks, performed with two null models: The *fixed* null model [60] that preserves exactly the network's degree sequence (see the FF model in Section 3.2.1), and the *probabilistic* null model [23] that, on average, preserves the observed total number of interactions and the network's degree sequence (see the PP model in Section 3.2.1 and [39] for implementation details. Black and white circles represent significant and non-significant correlations, respectively.

Source: Reprinted from [39].

effects of a localized perturbation in a networked population of springtail (*Folsomia candida*) microarthropods. They found that a modular topology can limit the impact of a localized perturbation. More research is needed to assess the impact of different network topologies on the spreading of a localized perturbation.

Modular-to-nested transition in communication dynamics. While several socio-economic networks exhibit nestedness (see Section 2), how such nested patterns emerge remains elusive. Borge-Holthoefer et al. [32] found that in communication networks, the emergence of collective attention on a given topic can manifest itself as a transition from a phase where the system exhibits both significant nestedness and modularity to a phase where the system only exhibits significant nestedness – such kind of evolution is referred to as *modular-to-nested transition* [32]. In other words, when collective attention narrows down to a single topic, the system abruptly transitions from a phase where the attention is dispersed across different sub-topics (modular structure) to a phase where all the attention is concentrated on a given topic (nested structure, no significant modularity).

Borge-Holthoefer et al. [32] uncovered this phenomenon by analyzing a two-month time-resolved user-meme network from Twitter that only includes memes related to the 2011 civil protests in Spain (see [32] for details on data collection). They found that in a first phase, both nestedness and modularity exhibit a growing trend. However, around the climax of the civil protests (May 15–17), they observed a transition point after which nestedness keeps increasing, whereas modularity severely decreases. Such a transition marks the beginning of a phase where network topology is radically different from the initial topology (Fig. 16). This result suggests that in communication networks, a nested architecture can emerge as a result of the emergence of a consensus on a given topic. A more recent study [141] also focused on temporal variations of the topology of a Twitter communication network, and found that centralization tends to increase when the discussion becomes more specialized.

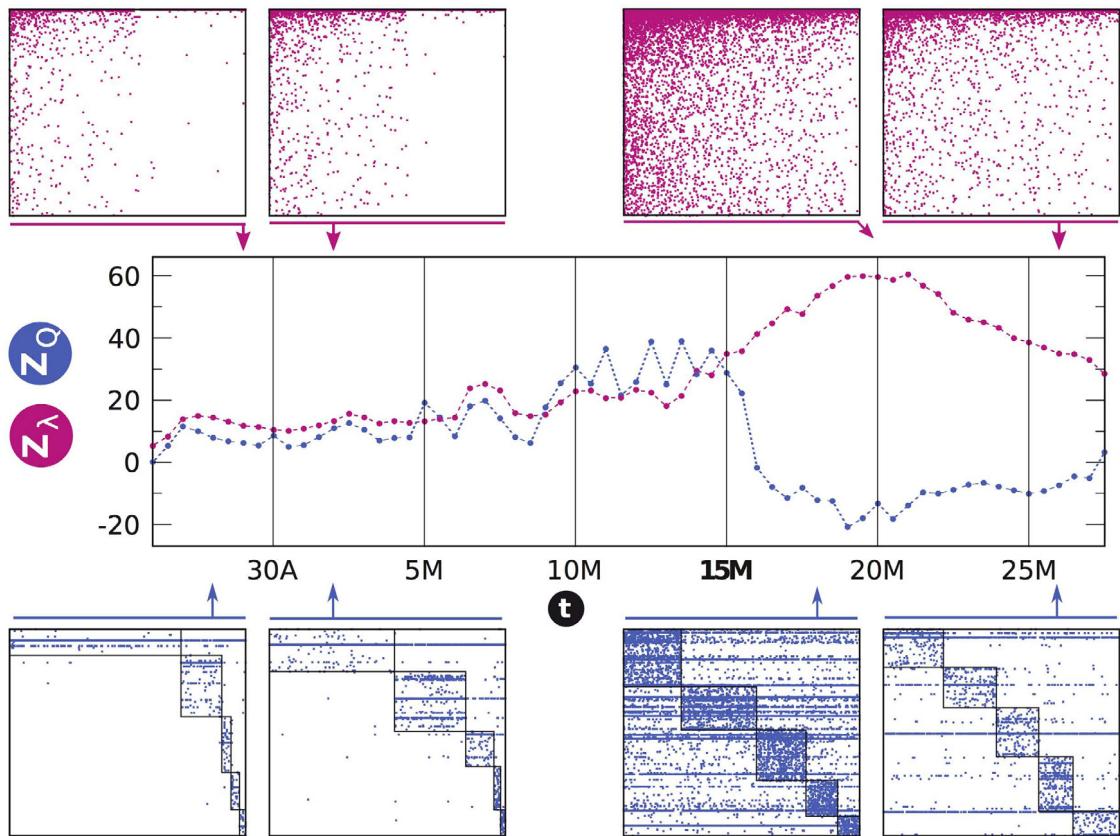


Fig. 16. Modular-to-nested transition in the bipartite user-meme network that only includes memes related to the 2011 civil protests in Spain. The central panel shows the simultaneous temporal evolution of nestedness z_λ and modularity z_Q over a two-month period (April–May 2011) – both quantities are z-scores obtained by comparing the observed level of nestedness and modularity, respectively, with a suitable null model. The other panels show the network's adjacency matrix when rows and columns are ordered in order of decreasing degree (top panels) and according to the detected modules (bottom panels). Importantly, around the 15 May 2011, we observe the abrupt modular-to-nested transition described in the main text.

Source: Reprinted from [32].

4.2. Nestedness and ecological properties

Given an observed structural pattern, it is important to assess how that pattern is related to other system-specific macroscopic properties. In ecology, an interaction network is embedded in an external environment; exploring how network nestedness (and modularity) is related to the macroecological environmental properties (such as precipitation, mean annual temperature, and temperature variability) has been the subject of various studies in ecology (see Section 4.2.1). The level of nestedness of spatial networks has been also related to their beta diversity [250], an important property that quantifies the relative magnitude of local and regional diversity (see Section 4.2.2).

Besides, from a complex systems perspective, an essential challenge is to explain how the observed structural pattern of interest arose from the interactions of its constituents [251–253]. To this aim, one can adopt two (possibly complementary) strategies. From a network science perspective [10,11], one can study dynamic mechanisms of network growth that seek to reproduce the observed structural patterns [13,25,36,254], and statistically validate mechanisms that explain the dynamics of the system [255]. This aspect is further developed in Section 5 where we introduce several distinct network formation mechanisms that aim to explain the observed level of nestedness in ecological and socio-economic systems.

From an ecological perspective, one seeks to explain how the observed individual interactions are mediated by species individual traits. Explaining what drives pairwise interactions can, in turn, explain why macroscopic structural patterns of interest have been found. Ecologists have identified species abundance [68,256], the existence of forbidden links [68,256], and phylogenetic signal [257] as factors that can, to some extent, explain the observed degrees of nestedness (see Section 4.2.3).

4.2.1. Nestedness and macroecological properties

We summarize here some of the results on the relation between nestedness and macroecological properties, i.e., ecological properties that characterize the system as a whole. Nestedness and modularity display linear relationships with

precipitation in plant-pollinator networks [258]. In a similar spirit, Schleuning et al. [259] focused on modularity, and found that modularity is negatively and positively correlated with mean annual temperature and temperature seasonality, respectively, in weighted seed-dispersal networks. Welti and Joern [260] examined 46 bipartite mutualistic networks and 22 bipartite trophic networks. They found that nestedness decreases with temperature variability between years. On the other hand, no significant evidence was found for an influence of environmental variables on the level of nestedness of trophic networks [260].

Dalsgaard et al. [261] studied how climate change affects nestedness and modularity of pollination networks. They found that the magnitude of this influence is different among islands and mainland regions. On the mainland, they found a positive association between nestedness and Quaternary climate change. The same was not found on islands, where Quaternary climate-change had no effects on nestedness. Sebastián-González et al. [262] found that nested structures tend to be located “in areas with a high degree of human impact, high temperature seasonality, low precipitation, and, especially on the mainland, high stability in precipitation”. Takemoto and Kajihara [263] confirmed that nestedness tends to increase with human impact, whereas modularity follows the opposite trend. Finally, we refer to [264] for a recent review on the variation of structural patterns of ecological networks along environmental gradients.

4.2.2. Spatial networks: Nestedness and beta diversity

Beta diversity is a key concept in biogeography. Broadly introduced by Whittaker [250] as “The extent of change in community composition, or degree of community differentiation, in relation to a complex-gradient of environment, or a pattern of environments”, beta diversity has been quantified through various metrics and its definition has been revisited several times. We refer to [265] for an exhaustive review. In this Section, we introduce the original definition of beta diversity provided by Whittaker [250] and discuss the relation between beta diversity and nestedness metrics. In addition, we will mention recent developments on the topic.

For a given geographical region composed of multiple sites, Whittaker’s β is defined as the ratio between the regional and local diversities. In formulas, the regional diversity (usually referred to as “gamma” diversity) can be quantified as the total number of species in the geographical region, $\gamma = N$. The local diversity (“alpha” diversity) can be quantified as the average number of species per geographic site, $\alpha = M^{-1} \sum_{\alpha=1}^M k_\alpha = \langle k^c \rangle$. Whittaker’s β is defined as the ratio between regional and local diversity: $\beta := \gamma/\alpha = N/\langle k^c \rangle = NM/E = \rho^{-1}$. Essentially, Whittaker’s β can be interpreted as a quantification of “how many times as rich in effective species the dataset is than one of its constituent compositional units” [265]. If one adopts Whittaker’s β to quantify beta diversity, the relation between nestedness and beta diversity is determined by the relation between the adopted nestedness metric and network density. For example, the nestedness metric by Wright and Reeves (Eq. (9)) is negatively correlated with β [159].

Beta diversity can be also quantified in terms of the dissimilarities between the species compositions of different sites. For a two-site region, one can define the dissimilarity between the two sites through the Sørensen index [266]

$$\beta_{sor} = \frac{b + c}{2a + b + c}, \quad (56)$$

where a denotes the number of species found in both sites; b denotes number of species that found in the first site but not in the second site; c denotes number of species that found in the first site but not in the second site. This definition can be readily generalized to multiple-site regions [267].

Importantly, ecologists have debated the factors determining the level of beta diversity of a given geographical area. Baselga [267] emphasized that beta diversity is modulated by two properties of the region: nestedness and turnover of species between sites. It is indeed intuitive that perfectly nested regions must exhibit some degree of diversity: in line with the definition of nestedness, some of the species that are present in richer sites cannot be found in poorer sites. At the same time, even a low-nestedness region can exhibit beta diversity if different species populate the different sites (“turnover of species”) – see Fig. 1 in [267] for an illustration.

Stimulated by these considerations, a question arises: how to disentangle the relative contributions of nestedness and species turnover to beta diversity? The Sørensen index defined by Eq. (56) is, in principle, affected by both factors. Following Baselga [267], one can consider a dissimilarity index that is only affected by species turnover and not by nestedness, like the Simpson index [268]. For a two-site region, the Simpson index β_{sim} is defined as [267]

$$\beta_{sim} = \frac{\min\{b, c\}}{a + \min\{b, c\}}. \quad (57)$$

If two sites are occupied by the same number of species, nestedness does not contribute to the region’s beta diversity. Therefore, the region’s beta diversity is entirely determined by species turnover, and $\beta_{sor} = \beta_{sim}$. Such equality does not hold, in general, for regions composed of two sites that have a different number of species. In that case, one can quantify the contribution of nestedness to beta diversity as $\beta_{nes} = \beta_{sor} - \beta_{sim}$, which implies that the overall diversity β_{sor} is modulated by the joint contribution of species turnover (β_{sim}) and nestedness (β_{nes}). By performing the subtraction and using Eqs. (56)–(57), one obtains [267]

$$\beta_{nes} = \beta_{sor} - \beta_{sim} = \frac{\max\{b, c\} - \min\{b, c\}}{2a + \max\{b, c\} + \min\{b, c\}}(1 - \beta_{sim}). \quad (58)$$

We refer to [267] for the generalization of these indexes to regions with multiple sites ($M > 2$), and to [267,269] for an analogous decomposition of the β diversity for β -diversity metrics different than the Sørensen index (e.g., Jaccard dissimilarity [269]). Beta-diversity metrics and their contributions from nestedness and species turnover can be found in the R package `betapart` [270].

4.2.3. Interaction networks: Nestedness and relative species abundance distribution

Ecologists have recognized since a long time that species abundance plays an important role in shaping the structure of mutualistic networks [8,9]. A natural question arises: can species abundance explain the observed levels of nestedness? The question has been investigated in the context of neutral theory [271], a theoretical framework that assumes that all species within a particular trophic level¹⁷ have the same chances of reproduction and death regardless of their peculiar traits [272]. In other words, according to neutral theory, all species are functionally equivalent: if an individual of a given species dies, it is replaced by a new individual, regardless of the species it belongs to. This assumption and neutral-theory stochastic models¹⁸ lead to expected relative species abundance (RSA) distributions [273]. By assuming that the species–species interaction probabilities only depend on species abundance, one can check whether the abundance distributions predicted by neutral theory can explain the observed degree of nestedness [68].

While such a parsimonious explanation of nestedness is tempting, it might neglect other important factors. It is known indeed that unobserved interactions are, in most cases, the consequence of biological constraints that make that interaction physically impossible – see [274] for an evaluation of the number of such “forbidden links” in interaction networks.¹⁹ Can we explain the observed nestedness through species abundance, without including forbidden links? If we include the forbidden links, can we achieve a more accurate explanation of the observed nestedness? Another caveat is necessary for the use of the word “explanation”. Even if species abundance alone was able to explain, in a statistical sense, species degree and network nestedness, it is still debatable whether generalists are generalists because they are more abundant or, by contrast, more abundant species are more abundant because they are more generalists [256].

With these caveats in mind, one can investigate whether the Relative Species Abundance (RSA) distribution predicted by neutral theory can explain the observed nestedness of mutualistic networks [68]. The Relative Species Abundance (RSA) distribution of a given community is denoted as the sequence $\{n_1, \dots, n_S\}$ of the species' population sizes, where S is the total number of species. To uncover the impact of the RSA distribution on nestedness, Krishna et al. [68] proceeded in two steps. First, they generated the plant and animal RSA distributions based on the theoretical prediction of neutral theory [273]. We denote the two distributions as $\{p_1, \dots, p_M\}$ and $\{a_1, \dots, a_N\}$, respectively, where M (N) denotes the total number of plant (animal) species. Second, they assumed that the probability P_{ia} that animal i interacts with plant α is proportional to their relative abundance: $P_{ia} = a_i p_\alpha / (\sum_j a_j \sum_\beta p_\beta)$. They built mutualistic networks – which they refer to as *neutral mutualistic networks* – by establishing a link between a plant and animal if a number of interactions above a given threshold was observed. Finally, they investigated whether the resulting networks exhibit a nested structure.

Besides, they aimed at establishing whether the RSA distribution could explain the nestedness observed in mutualistic networks. To this end, they observed that if a given factor explains nestedness, it should produce a ranking of the nodes that well approximates the ranking by the nestedness temperature calculator, which is aimed at maximizing the degree of nestedness (see Sections 3.3 and 3.1.2). Based on this observation, the problem of assessing whether the RSA distribution explains nestedness reduces to comparing the temperature T_{ab} of the matrix where species are ranked by their abundance with the temperature T of the matrix that is produced by the nestedness temperature calculator.

The results of their investigation are two-fold. First, they found that neutral mutualistic networks are almost perfectly nested. Probabilistic animal–plant interactions consistent with neutral theory [273] are, therefore, among the possible mechanisms that lead to the emergence of nestedness. Second, they found that in real mutualistic networks, the RSA distribution can explain only part of the observed nestedness, with values of the fraction of explained nestedness $(100 - T_{ab})/(100 - T)$ around 0.6 – 0.7 [68]. Importantly, they found that the agreement between expected and observed nestedness can be substantially improved by accounting for the forbidden links.

Santamaría and Rodríguez-Gironés [256] found that a log-normal neutral model of interactions (i.e., a model that produces a log-normal RSA distribution) fits well the observed nestedness in empirical pollinator–plant networks. However, they pointed out that the log-normal neutral model should be rejected as a parsimonious explanation of nestedness because of three factors [256]: (1) the fact that the model fits the empirical nestedness does not tell us anything about the causal direction of the abundance-degree relationship: it is not excluded that more abundant species might be more abundant because they are generalists, and not the other way around; (2) the RSA distribution might explain nestedness under the assumption that the species–species interaction probability is proportional to species abundance; however, this assumption is not supported by empirical studies [276]; (3) As we pointed out above, several studies [274,277] have shown that forbidden links play a critical role in shaping the topology of interaction networks.

¹⁷ Trophic level refers to the set of all species that belong to the same level in the food chain [271]. Neutral theory is limited to communities of species that belong to one specific trophic level.

¹⁸ A detailed description of the models and implications of neutral theories goes out of the scope of the present review, but it can be found in the review article by Azale et al. [271].

¹⁹ Among the unobserved links, forbidden links are fundamentally different from those links (“missing links” [274] or “neutral forbidden links” [275]) that were not recorded simply because we did not observe the system over a long enough time span.

Hence, while of undeniable relevance, the connection between nestedness, species abundance, and forbidden links alone does not allow us to infer the generative mechanisms that have led to those macroscopic properties.

Beyond relative species abundance distribution, Rezende et al. [278] aimed to detect the presence of a phylogenetic signal for plants and animals assemblages for 36 plant-pollinator and 23 plant-frugivore mutualistic networks. To detect phylogenetic signal, they compared the properties of the empirical networks with those numerically generated based on the structure of the species phylogenetic history [257,278]. They found a fundamental difference between species degree (i.e., number of interactors per species) and species strength (i.e., number of recorded interactions): differently from species strength, species degree exhibits a strong phylogenetic signal [9,278]. Besides, they found that the distance between two given species in the interaction network is significantly correlated with their distance in the phylogeny: for most communities, phylogenetically-similar species tend to play similar roles in the interaction network. In a different article, Rezende et al. [257] found that phylogenetic effects can significantly contribute to nestedness. Overall, information on relative abundance and phenology can be sufficient to predict the level of nestedness of a network, yet it cannot predict the detailed network structure [279].

4.3. Nestedness and economic properties

Some scholars have argued that the ultimate goal of understanding a complex system is to accurately predict its future behavior [280]. Predicting the future evolution of a complex system can be difficult due to various factors, including chaotic behavior, our incomplete knowledge of the fundamental laws that describe the system's behavior, our lack of knowledge of the system's state vector – we refer to [281] for a survey. The network science literature has designed a wide spectrum of techniques either to predict the future links that will appear in a network, or to infer the connections that are missing in noisy data (see [280,282,283] for recent reviews on the topic). The nested structure of socio-economic systems has motivated scholars to leverage such structure to predict the future behavior of the system, at the level of both individual links and nodes. These studies aimed to predict the future success of firms [28] (Section 4.3.1) and countries [121,122] (Section 4.3.3), and to predict future links in economic networks [29] (Section 4.3.2). Besides, in a similar spirit as neutral models in ecology (see Section 4.2.3), scholars have attempted to explain the observed level of nestedness in economic networks based on hidden capabilities that drive the formation of pairwise interactions (see Section 4.3.4).

4.3.1. Individual nestedness contributions and the survival of firms

As nestedness arises as a result of node-level interactions, it is natural to ask whether some nodes give a larger contribution to the nested structure of a given system. Should this be the case, one can further ask (1) whether the strongest contributors to nestedness play a more significant role to network persistence (measured as the fraction of initial nodes that did not become extinct by the end of population dynamics simulations [28]) than weaker contributors; (2) whether stronger contributors experience a larger or smaller probability of becoming extinct.

Saavedra et al. [28] quantified the individual contribution to network nestedness from each individual node. Given a node i , its contribution C_i to nestedness is determined by comparing the observed nestedness with the nestedness of randomized networks obtained by only randomizing node i 's links. More specifically, $C_i = (\mathcal{N} - \mu_i(\mathcal{N})) / \sigma_i(\mathcal{N})$, where $\mu_i(\mathcal{N})$ and $\sigma_i(\mathcal{N})$ represent the mean and the standard deviation, respectively, of nestedness \mathcal{N} over the networks where node i 's links have been randomized.

Saavedra et al. [28] found that both in ecological mutualistic networks and in economic manufacturer-contractor networks, individual contributions to nestedness are heterogeneously distributed, and some nodes give a significantly larger contribution than the others. Through simulations on mutualistic networks based on a model of mutualistic interactions,²⁰ they found that the strongest contributors to nestedness are also those whose extinction has the largest impact on network persistence.²¹ The results of numerical simulations indicate that, ironically, the strongest contributors are also the most vulnerable nodes to extinction [28].

While the analysis of the persistence and stability of ecological communities almost invariably requires a population dynamics model, datasets on social and economic activity often feature fine-grained temporal information, which can relieve us from the need of numerical simulations. In particular, one can translate the questions above to economic systems: have the firms with the strongest contributions to nestedness the highest risk to fail? Saavedra et al. [28] addressed this question by analyzing a manufacturer-contractor bipartite network (see Section 2.3.2) over a 15-years time span. In agreement with their model-based findings on mutualistic networks, they found that for both manufacturers and contractors, the strongest contributors to nestedness are also the firms that are less likely to survive.

The work by Saavedra et al. [28] highlights at least two important factors. First, the similarity of the findings in ecological and economical networks point to the generality of the role of nestedness individual contributions for node survival. Second, ecologists often lack time-stamped data over a long enough time span. As a result, works on the implications of network topology on systemic stability are typically based on population dynamics and dynamical-system stability analysis (see Section 6). On the other hand, socio-economic data are often time-stamped. Socio-economic data might, therefore, allow us to validate mechanisms on the emergence of nestedness and its implications, which might give insights that are also relevant to ecologists.

²⁰ We refer to [28] for details. The model is essentially the same that has been applied in other works [38,50] to uncover the impact of nestedness on systemic feasibility and dynamical stability – see Section 6.2.3.

²¹ A recent study [71] indicates that non-linear iterative algorithms perform better than individual nestedness contributions in identifying the most important and vulnerable network nodes under targeted attack. See Section 6.1.2 for a detailed presentation.

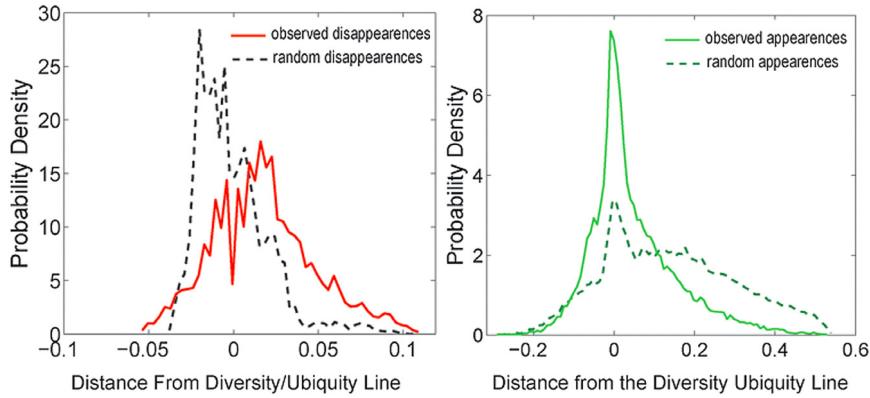


Fig. 17. Fraction of appearances and disappearances as a function of the distance from the diversity–ubiquity line for the country–product network. Observed appearances (disappearances) tend to be closer to (farther away from) the diversity–ubiquity line than expected by chance. Source: Adapted from [29].

4.3.2. Nestedness and prediction of economic links

The study by Saavedra et al. [28] focused on the prediction of the survival of firms. At a finer scale, one might attempt to leverage the nested structure to predict the future links of a system. The argument is the following: suppose that we are confident that a given system exhibits a significantly nested structure. If we found a link in a matrix region where there are only a few other links, we might argue that this link constitutes an outlier in the overall systemic structure and, as a consequence, it is more likely to disappear in the future. In an analogous way, if a link is missing in a dense region of the adjacency matrix, it might be more likely that link will appear in the future.

Building on a previous work in ecology [284], Bustos et al. [29] tested this idea in both the country–product bipartite network over the 1985–2009 period, and in a Chilean firm-location bipartite network over the 2005–2008 period. For both networks, they found that the level of nestedness was relatively stable over time and statistically significant with respect to the Proportional–Proportional model by [23] (see Section 3.2.1) and a dynamic null model (see Figs. 1f-g in [29]). Besides, they found that the coefficients of the logistic model [29]

$$M_{i\alpha}(t) = a k_i(t) + b k_\alpha(t) + c k_i(t) k_\alpha(t) + \epsilon_{cp}(t) \quad (59)$$

are all highly significant: the interaction term $c k_i(t) k_\alpha(t)$ is essential to explain the presence of links. This interaction term is compatible with the nested structure of the network, as it penalizes links between two nodes with low degree. The observed temporal persistence of nestedness, together with the explanatory power of the model (59), suggests that nestedness can be used to predict the future links in both systems.

To test this hypothesis, Bustos et al. [29] compared, for each pair of years, the ability of their model to predict the appearance and disappearance of links.²² The model produces an Area Under the Curve (AUC)²³ substantially larger than 0.5 for both appearances and disappearances, which implies better predictability than random predictions. Interestingly, it predicts disappearances significantly better than appearances (average AUC equal to 0.81 against 0.62, see Fig. 3c in [29]). Besides, they studied the structural position of appearances and disappearances in the adjacency matrix. In particular, they measured their distance from the diversification–ubiquity line, which denotes the shape of a perfectly nested matrix with the same number of edges as the matrix in exam. They found that the observed appearances tend to lie significantly closer to the diversification–ubiquity line than expected by chance, whereas observed disappearances tend to lie outside of the perfectly nested shape (see Fig. 17). The nested structure of the network is therefore highly informative on future changes in the system.

More recently, Medo et al. [285] aimed to leverage the nested structure of World Trade and mutualistic networks to reconstruct missing links in these systems. More specifically, they considered the number of violations of the nestedness condition as defined by the following equation [43,285]

$$V := (1 - \Theta(k_j - k_i)) \sum_{\alpha} A_{j\alpha} (1 - A_{i\alpha}). \quad (60)$$

Based on this quantity, they defined the link-prediction score of a link (i, α) , $s_{i\alpha}$, as the difference between the number of violations of the nestedness condition after (i, α) is added to the network, $V^{(i, \alpha)}$, and the number of violations of the

²² Bustos et al. [29] defined an appearance of a link (i, α) as an increase in exports per capita from less than 5% of the world average, for five consecutive years, to more than 25%, for at least five years. They defined the disappearance of a link (i, α) as the decrease in exports per capita of a country from 25% or more than the world average to 5% or less, for at least five years.

²³ The AUC is one of the most used metrics to benchmark prediction accuracy. We refer to [29,282] for its definition.

original network, V : $s_{i\alpha} = V^{(i,\alpha)} - V$. They validated this index in both synthetic networks with a tunable degree of nestedness and empirical networks, finding that the resulting link-prediction method can outperform some of the state-of-the-art link prediction techniques for both highly-nested synthetic networks and some of the empirical networks that they analyzed.

Beyond the nestedness-based method by Bustos et al. [29] and Medo et al. [285], scholars attempted to predict future links in World Trade based on country-level development paths in the product space [286,287], and on recommender systems [288]. We point out that a comprehensive evaluation of different methods for link prediction in World Trade is currently missing. Such comparative analysis would be particularly relevant as World Trade data are typically noisy due to diverse factors [45,198], which can affect the reliability of structural metrics of country and product importance [72,115,198] and the resulting economic development forecasts [45].

4.3.3. Nestedness maximization and the future development of countries

As we have discussed in Section 3.3.2, ranking algorithms that produce nested adjacency matrices (such as the fitness-complexity algorithm [24]) have been extensively applied to world trade networks, aiming to quantify the countries' competitiveness and the products' level of sophistication. The fitness score of a country can be interpreted as a proxy for its number of available "capabilities", to be interpreted as the building blocks needed to produce and export different types of products [110] (see Sections 3.3.2 and 4.3.4). Besides, the countries' fitness score is positively correlated with their gross domestic product per capita (GDPpc), yet this correlation is far from one [24,121].

Importantly, the deviations from the linear-regressed trend are informative about the future development of countries [121]. Suppose indeed that a country has a relatively ordinary GDPpc, and a relatively large fitness score. This means that while the level of development of that country is not yet comparable with that of the leading economies, the country possesses a wide set of capabilities which might allow it to economically grow in the future. The recent history has seen one such example: China. In 1995, China had a substantially larger fitness than countries with a similar GDPpc and, in the subsequent years, China's GDPpc has experienced a drastic increase [193].

The hypothesis that a country's high fitness score to economic growth in the future has been first tested by Cristelli et al. [121]. They analyzed 16 years of World Trade (1995–2010), and sought to predict the future development of the countries based on their position in the GDPpc-fitness plane (see Fig. 18). Of course, given the inherent complexity of the world production, it is not obvious that the history from the past is informative about the future of the system. If the current fitness of a country is informative about its future development, we would observe highly-predictable regions in the fitness–GDPpc plane. That would signal some degree of "stationarity" in those regions, and the possibility to attempt to predict the future [122].

Borrowing from the methodologies from the dynamical systems literature [289], Cristelli et al. [121] built the Selective Predictability Scheme (SPS) which analyzes the past history of the system to evaluate, for each region of the plane, its level of predictability²⁴ and the most likely future position in the plane for countries that are located in that region. Intriguingly, there turn out to be both highly-predictable and low-predictability regions in the fitness–GDPpc plane [121]. Importantly, the predictability of a region is an increasing function of fitness, and high-fitness countries tend indeed to grow their GDPpc in the future [121,122]. When visualizing the dynamics of countries through animations, one can detect by eye the regular upward movement of the countries in the high-fitness region [290].

A recent work by the World Bank [122] systematically compared the predictions of the SPS based on country fitness with those by the International Monetary Fund (IMF). They found that, when one only considers countries that are in the predictable region of the fitness–GDPpc plane, the SPS predictions significantly outperform the IMF predictions. The performance of the SPS and IMF are comparable when one considers all the countries. With respect to the IMF growth model which builds upon many variables, the SPS method is very parsimonious, as it only involves one parameter: country fitness. In a subsequent work, Tacchella et al. [45] found that proper data sanitation can substantially increase the volume of the predictable region; on the cleaned data, the predictions by the SPS method over a five-year time horizon outperform those by the IMF even when one considers both high- and low-fitness countries. Besides, the predictions by the SPS can be combined with those by the IMF to substantially improve the GDPpc forecasts. Myanmar, Cambodia, and India turned out to be the three countries whose GDPpc's forecast benefited most from the combination of SPS and IMF results [45].

These works have inspired other scholars [291–293] to investigate the relation between a country's fitness and its future economic development. To conclude, we mention two additional aspects of this framework. First, in a similar way, one can study the dynamics of products on a plane formed by the score of the product by the fitness-complexity algorithm and a suitable indicator of product added value [294]. Second, one can build score-GDPpc planes for the scores produced by other structural metrics based on the bipartite country–product network. In particular, Liao et al. [291] measured the overall predictability²⁵ of the score-GDPpc plane for the country-level scores by three different network-based ranking metrics: degree, fitness, and the Economic Complexity Index (ECI) [295] – see Fig. 15 in [291] for an illustration. They

²⁴ The predictability of a region \mathcal{R} over a t -years time window is essentially the heterogeneity of the region $\mathcal{R}(t + \Delta)$ where countries that were in region \mathcal{R} at time t are located at time $t + \Delta t$. In practice, it is convenient to split the plane into rectangular boxes and study the predictability of each of the resulting boxes [121].

²⁵ The overall predictability of the plane is obtained by aggregating the predictability of its regions, and weighting each region by the number of observed events – see Section 7.2 in [291] for details.

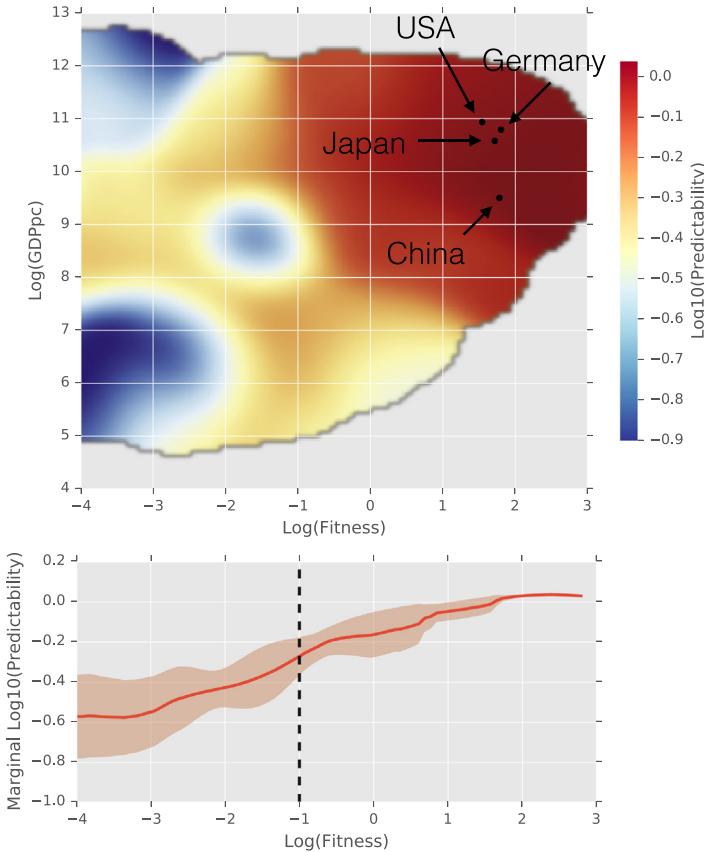


Fig. 18. Economic-complexity predictions of the national Gross Domestic Product per capita (GDPpc): Level of predictability in the fitness–GDPpc plane (top panel), and as a function of fitness (bottom panel). The GDPpc of high-fitness countries is significantly more predictable than that of low-fitness countries.

Source: Reprinted from [122].

found that country fitness as quantified by the fitness–complexity algorithm leads to the largest overall predictability, whereas no regular dynamics emerges in the ECI–GDPpc plane. The obtained results suggest a link between nestedness maximization (see Section 3.3) and economic predictions: rankings of countries that produce highly-nested matrices, such that the one by fitness, can be highly informative for the countries’ future development.

4.3.4. Nestedness and hidden economic capabilities

In the economic complexity approach to World Trade analysis [24,110,121], the observed bipartite country–product export network is interpreted as the result of a projection of a tripartite country–capability–product network where countries are connected with the capabilities they possess, and products are connected with the capabilities they require in order to be produced (see Fig. 9). Hausmann and Hidalgo [52] leveraged this intuition to build a model of the World Trade where a given country exports a given product if and only if the country possesses all the capabilities required to produce that product. In practice, this model can be considered as a “neutral model” of development [29,52] in a similar spirit to neutral models in ecology (see Section 4.2.3): the interactions driven by a hidden property – species abundance and country capabilities in ecological and trade networks, respectively – might explain the network’s structural patterns and, in particular, its nestedness.

Formally, let us denote by $\kappa \in \{1, \dots, K\}$ the available capabilities. We denote by A the bipartite country–product adjacency matrix, by C the country–capability matrix ($C_{ik} = 1$ if country i possesses capability κ), by P the product–capability matrix ($P_{\alpha\kappa} = 1$ if product α requires capability κ in order to be made). With this notation, $A_{i\alpha} = 1$ if and only if $\sum_{\kappa} C_{ik} P_{\alpha\kappa} = \sum_{\kappa} P_{\alpha\kappa}$, i.e., if country i possesses all the capabilities necessary to produce α . Therefore, the model assumes that the availability of capabilities is the only factor that determines the level of diversification of each country: each country will produce all the products for which that country has the required capabilities. While in contrast with standard economic theories that prescribe country specialization as a determinant for wealth [118], this assumption well matches the empirical nestedness of the country–product matrix [24,29].

The matrices that involve the capabilities, C and P , are unobservable; nevertheless, one can formulate a probabilistic rule to generate their elements, and verify a posteriori the agreement between the generated country–product matrices

and the observed ones. Hidalgo and Hausmann [110] assumed that country i possesses capability κ with probability r , whereas it does not possess it with probability $1 - r$; analogously, product α requires capability κ with probability q , whereas it does not require it with probability $1 - q$. Hence, each country i ends up with a given number of available capabilities $c_i = \sum_{\kappa} C_{ik}$, and each product with a given number of required capabilities $p_{\alpha} = \sum_{\kappa} P_{\alpha\kappa}$. The model has three parameters: r , q , C . Nevertheless, to calibrate the model to a real system, only two of these parameters are free, whereas the third one is fixed by the condition that the average degree of the model-generated networks matches that of the original network.

It is natural to investigate how countries' diversification and products' ubiquity are related to their number of available and required capabilities, respectively. The average number of products exported by country i (or i 's average diversification), \bar{k}_i , can be written in terms of i 's number of available capabilities, c_i , as

$$\overline{k_i(c_i)} = \sum_{x=1}^C \mathbb{P}[c_i \rightarrow x] M(x), \quad (61)$$

where $\mathbb{P}[c_i \rightarrow x]$ denotes the probability that a country with c_i capabilities exports a product that requires x capabilities, whereas $M(x)$ denotes the number of products that require x capabilities. Given that each capability is extracted independently, the probability $\mathbb{P}[c_i \rightarrow x]$ is simply given by $\mathbb{P}[c_i \rightarrow x] = (c_i/C)^x$. The fraction of products that require x capabilities is given by the binomial distribution:

$$\frac{M(x)}{M} = \binom{C}{x} q^x (1-q)^{C-x}, \quad (62)$$

which is why the model was referred to as *binomial model* by Hausmann and Hidalgo [52]. Therefore, we obtain

$$\overline{k_i(c_i)} = M \sum_{x=0}^C \left(\frac{c_i}{C}\right)^x \binom{C}{x} q^x (1-q)^{C-x}. \quad (63)$$

By using the binomial series, we can approximate the previous expression as

$$\overline{k_i(c_i)} \simeq M \left(q \frac{c_i}{C} + 1 - q \right)^C. \quad (64)$$

In a similar way, one can obtain expressions for product ubiquity as a decreasing function of the number of required capabilities, and the expected functional forms of the country diversification and product ubiquity distributions,

While this model produces country–product networks that are significantly more nested than expected by chance, it cannot generate networks with a degree of nestedness comparable to the observed one in World Trade data [29]. To overcome this limitation, Bustos et al. [29] introduced an alternative capability-based model where the number of capabilities each country has is a random number extracted from the uniform distribution in $[0, R]$, whereas the number of capabilities required by each product is a random number extracted from the uniform distribution in $[0, Q]$. Because of such uniform distributions, the model is referred to as the *uniform model* by Bustos et al. [29]. The model has again two free parameters: R and Q . The probability that country i has capability κ is given by $P_{ik} = \min\{1, R i/c_i\}$. The uniform model generates bipartite adjacency matrices that are significantly more nested than those generated by the binomial model, and the level of nestedness of the generated matrices is comparable with that observed in real country–product matrices [29].

An alternative capability-based model was proposed by Tacchella et al. [296], who assumed that capabilities have a different “usefulness”, defined as the number of products that require that capability. The possibility of varying the usefulness of individual capabilities makes such model more general than the binomial model which assumes that the number of capability required by each product follows a binomial distribution. A power-law distribution of usefulness can be used to generate networks with levels of nestedness comparable with those observed in real data [296], which suggests that the most diversified countries might be those endowed with the most “useful” capabilities.

5. Emergence of nestedness: Rewiring and formation mechanisms

Network structural properties are of interest not only because they deepen our understanding of the organization of complex systems, but also because distinct macroscopic patterns result from different microscopic mechanisms of network growth. Candidate mechanisms to describe the evolution of the system might need to be ruled out if they lead to network topologies that mismatch those observed in real data. Following the seminal work by Barabási and Albert [13], explaining the scale-free degree distribution of many real systems has received an enormous amount of attention in the network science literature [297]. Models with a node-level hidden variable (typically referred to as *fitness* [298]) provide an explanation for the scale-free structure of real networks that does not resort to preferential attachment mechanisms [298]. Importantly, a fitness model with a threshold linkage rule [53,54] can generate perfectly nested networks with various degree distributions (see Section 5.1).

In parallel, ecologists have devoted a massive effort to understand how nested networks can emerge in interaction networks. Mechanisms that have been proposed include (but are not limited to) competition load minimization [50],

fitness optimization [33,299], trait-matching [27], speciation and divergence [34], invasion dynamics [35]. Some of these mechanisms build on optimization problems where species aim to maximize their own fitness [33,299] or the community-level fitness [33], whereas other mechanisms [34,35] do not involve optimization. Section 5.2 surveys the mechanisms for the emergence of nestedness in socio-economic networks.

More recently, scholars have started investigating the emergence of nestedness in socio-economic networks as well. In unipartite socio-economic network, this effort has revealed that nestedness can emerge as a result of a process where individual actors strive to maximize the centrality of their interaction partners, i.e., their social status [25,36,81], through a dynamical process that has been referred to as “social climbing game” [81]. From this standpoint, highly nested structures correspond to highly hierarchical, “low-temperature” structures social mobility faces high energetic costs [81]. In bipartite country–product export networks, capability-based mechanisms [29,52] can generate networks with a level of nestedness compatible with the observed one. Nestedness in World Trade has been also explained in terms of network formation mechanisms that feature both innovation and specialization in the countries’ development process [51]. Section 5.3 surveys the mechanisms for the emergence of nestedness in socio-economic networks.

5.1. Graph-theoretic mechanisms

Without even considering ecological or socio-economic processes, perfectly nested topologies can simply arise from a fitness model based on a threshold linkage rule [53,54,298]. In the following, we first review the threshold-based fitness model (Section 5.1.1), and then describe how threshold models can be used to generate nested networks with various degree distributions (Section 5.1.2).

5.1.1. Threshold model

The threshold model [298] is perhaps one of the simplest generative models for threshold networks, i.e., perfectly nested networks (see Section 2.1). Fitness models [298] are heterogeneous random graph models where each node i is endowed with one parameter η_i , referred to as fitness. The fitness value of each node is randomly drawn from some probability distribution $P(\eta)$; for each pair of nodes i and j , the probability P_{ij} that the two nodes are connected by a link is only a function of η_i and η_j : $P_{ij} = P(\eta_i, \eta_j)$. The threshold model [298] assumes that $P(\eta_i, \eta_j) = \Theta(\eta_i + \eta_j - \zeta)$, i.e., i and j are connected if and only if the sum of their fitness values is larger than a threshold value ζ .

Caldarelli et al. [298] studied the threshold model with an exponential fitness distribution $P(\eta) = \exp(-\eta)$, with $\eta \in [0, \infty)$. They found that the model generates networks with a power-law degree distribution. Back in the early 2000s, in the rising wave of interest in scale-free networks spurred by the seminal work by Barabási and Albert [13], this result was intriguing: to generate a scale-free network, neither growth nor preferential attachment are necessary elements. By adopting a generating-function approach, Boguñá and Pastor-Satorras [300] performed an extensive analytic investigation of the threshold model with exponential fitness distribution. They found that the degree distribution essentially decays as k^{-2} , yet there is an accumulation point at $k = N$ which corresponds to a small fraction of generalist nodes that are connected to all the other nodes in the network. Besides, they found that the average degree of the neighbors of nodes of degree k tends to decay as k^{-1} , which means that the network is strongly disassortative. As we have seen in Section 4.1.2, disassortativity is one of the fingerprints of nestedness.

Masuda et al. [53] emphasized that the threshold model generates threshold graphs (i.e., nested networks), and drew analogies with other threshold social phenomena that are typically investigated by means of complex contagion models [301–303]. They considered various fitness distributions (referred to as “weight” distributions in [53]), and they were interested in determining which fitness distributions can explain the heavy-tailed degree distributions often found in real networks. While the power-law degree distribution stemming from an exponential fitness distribution was already documented by Boguñá and Pastor-Satorras [300], Masuda et al. [53] found that power-law tails of the type k^{-2} can also arise from a logistic, Gaussian, and a Pareto fitness distribution. They pointed out that power-laws with exponent $\gamma = -2$ might be a general feature of thresholding models, as opposed to power-laws with exponent $\gamma = -3$ that emerge from the Barabási–Albert preferential attachment model [13].

5.1.2. Growing perfectly nested networks with arbitrary degree distribution

As the threshold model [298] can generate perfectly nested networks for different choices of the fitness distribution [53], one might wonder whether it would be possible to generate perfectly nested networks with arbitrary network topology. Hagberg et al. [54] provided a recipe to do that, approximately, for large enough networks. To this end, they introduced a creation-sequence mechanism for the generation of a threshold network, i.e., a perfectly nested network. The mechanism is simple: before creating the network, one partitions the nodes into the dominant and the independent set (see Section 2.1). At each time step, a new node is added to the network. If the newcomer is a dominant node, it connects to all the preexisting nodes. If the newcomer is an independent node, it remains isolated. It is evident that this growth mechanism generates a perfectly nested structure. Therefore, if we define a node-level binary variable x_i which is equal to one and zero for dominant and independent nodes, respectively, the growth of the perfectly nested network can be encoded in a creation sequence $S = \{x_i\}$. If we consider networks composed of five nodes, a star is represented by the sequence $S = 00001$, whereas the complete graph by the sequence $S = 11111$.

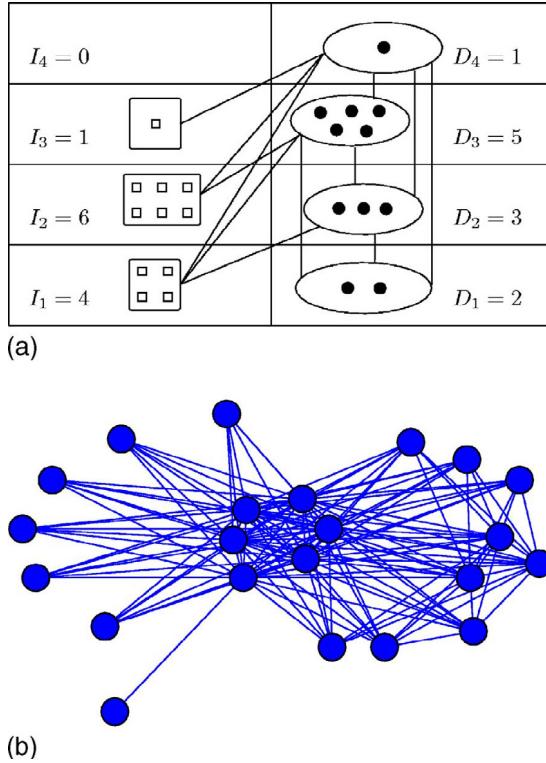


Fig. 19. Creation-sequence procedure to generate perfectly nested networks. From bottom to top in panel (a), one starts from $D_1 = 2$ dominant nodes, and subsequently adds $I_1 = 4$ independent nodes, D_2 dominant nodes, and so forth. When introduced into the system, independent nodes stay isolated, whereas dominant nodes connect to all the preexisting nodes. This results in a perfectly nested structure (panel (b)).
Source: Reprinted from [54].

An illustration of this growth mechanism is provided in Fig. 19. Interestingly, one can also show that this procedure is equivalent to the threshold model introduced in Section 5.1. This means that, given the set $\{\eta_i\}$ of nodes' fitness values and the threshold value ζ , one can obtain a unique creation sequence that generates the corresponding perfectly nested network. At the same time, given a creation sequence \mathcal{S} , one can find a unique set of fitness values and a threshold value that produce the corresponding perfectly nested network. We refer the interested reader to [54] for the proof of the equivalence.

By exploiting the growth dynamics presented above, we can generate perfectly nested structures with arbitrary degree distribution. To do that, the independent nodes are used to enforce the target degree distribution; indeed, given a creation sequence, the degree of a given independent node i is simply equal to the number of dominant nodes that are introduced after i . The dominant nodes turn out to have a large degree and they do not respect the pre-imposed degree sequence; however, their impact on the overall degree sequence is relatively small for sufficiently large networks.

5.2. Ecological mechanisms

The reasons behind the emergence of nestedness in ecological networks have been widely debated. Some works have interpreted the observed levels of nestedness in empirical networks as the potential result of optimization processes where species aim to maximize either their own fitness [33,80,299] or the community's fitness [33]. In this spirit, Medan et al. [80] found that nestedness can emerge when species switch their interaction in the attempt to increase the centrality of their neighbors (see Section 5.2.1). Saavedra et al. [27] found that a parsimonious bipartite-network formation model that features simple cooperation rules between the two classes of nodes can generate networks that exhibit levels of nestedness comparable to the empirically observed ones (see Section 5.2.2). Bastolla et al. [50] mentioned that if a species entering a community tends to minimize its competition load when establishing links, a nested structure would emerge. Zhang et al. [299] found that an interaction switch aimed at replacing the least profitable interactions (in terms of per-capita abundance gain) with more profitable ones can lead to nested structures comparable with the empirical ones. Suweis et al. [33] found that nestedness could emerge as a result of an evolutionary process where species aim to maximize their own fitness or the community's fitness (see Section 5.2.3).

Other works [34,35] have focused on nestedness and other structural properties as consequences of the generative rules of the system, without any optimization mechanism involved. These works pointed out that nestedness can be

observed in networks generated with selective mechanisms that do not act to adapt the species to their local environment. From this viewpoint, nestedness and other structural patterns might be viewed as analogous to spandrels in architecture: elements that might be perceived as functional key components of the system, yet they are mostly the byproduct of other structural components [304]. In this spirit, Valverde et al. [34] showed that nestedness can emerge as a result of a stochastic dynamics with speciation and divergence (see Section 5.2.4). Maynard et al. [35] pointed out that the presence (or absence) of nestedness might be the consequence of the system's specific assemblage rules, and seemingly minor differences in the way the network is assembled can lead to drastically different topologies (see Section 5.2.5).

In the following, we detail some of the above-mentioned mechanisms. We stress that the property alone that these mechanisms theoretically lead to a nested topology does not guarantee their empirical relevance. For example, even if a nested topology allows species to minimize their competition load [50] and maximize their fitness [33] according to analytic results on population-dynamics models, we have no guarantee that the observed structure of interaction networks has emerged for this reason. In other words, from the property that a topology (in our case, a nested topology) performs specific functions (in our case, minimization of competition loads and maximization of species' fitness), we cannot infer that the empirically observed structures were "selected" in order to perform those functions. A similar argument holds for mechanisms that feature no optimization of species-level or community-level properties [35,304]: the fact alone that a given mechanism can theoretically lead to nestedness does not imply that mechanism acted to shape the observed empirical networks.

5.2.1. Self-organizing network model

The self-organizing network model (SNM) introduced by Medan et al. [80] is one of the simplest models that lead to the emergence of nestedness. The model aims to explain the nested topology of ecological networks through a rewiring mechanism where the nodes tend to connect to central nodes in the system. We shall see that rewiring models based on a similar idea have been also applied to social systems [36,81] (see Section 5.3.3), which points out nodes' effort to connect to central nodes as a possible general mechanism behind the emergence of nestedness in diverse systems.

The SNM takes as input the number of species S , the total number of links E , and a set of forbidden links. One starts by distributing the E links across randomly-selected pairs of nodes. Starting from such random topology, we sequentially perform two iterative steps: (1) we select a row/column of the adjacency matrix, and we randomly pick a 1 and a 0 from the selected row/column. (2) We swap the two elements if and only if the following three conditions are met: the degree of the new partner is larger than the degree of the previous partner; the swap does not cause the previous partner to become extinct; the new link is not a forbidden link.

After a large number of iterations, this process leads to a perfectly nested structure, except when some pathological conditions occur.²⁶ As real networks are rarely perfectly nested, one can stop the iterations as soon as the network reaches the same level of nestedness as observed in the real network we aim to model. Interestingly, the model leads to truncated power-law degree distributions similar to those observed in real networks, regardless of the fraction of forbidden links [80]. The SNM was used by Burgos et al. [305] to assess the impact of nestedness on systemic robustness against node removal — see Section 6.1. A generalization of the SNM to incorporate phylogenetic data was introduced by Perazzo et al. [306] to uncover the influence of phylogenetic effects on nestedness.

5.2.2. Bipartite cooperation: Specialization and interaction

One of the common key aspects of both mutualistic ecological networks and manufacturer-contractor systems is that nodes benefit from interacting with each other. Motivated by this observation, Saavedra et al. [27] introduced a parsimonious network formation model based on bipartite cooperation. They used the model to explain the structural properties of both animal-plant mutualistic networks and of manufacturer-contractor networks; to fix ideas, in the following, we refer to the two classes of nodes as animals and plants, yet the model equally applies to bipartite networks of manufacturers and contractors. The model features two mechanisms:

- *Specialization* determines the number of partners, k_α , a given plant α interacts with. It is determined by

$$k_\alpha = 1 + \left\lfloor (L - M) \frac{r_\alpha \lambda_\alpha}{\sum_\beta r_\beta \lambda_\beta} \right\rfloor, \quad (65)$$

where r_α is the reward trait associated with interactions with plant α (randomly extracted from the uniform distribution in $[0, 1]$) and λ_α is a parameter (randomly extracted from an exponential distribution) that accounts for exogenous factors such as geographic effects and population diversity.

- *Interaction* determines which animals interact with each plant. Each animal i is endowed with a foraging trait f_i randomly extracted from the uniform distribution in $[0, 1]$. Plants are sorted in order of increasing reward trait r , whereas animals are sorted in order of decreasing foraging trait f . Besides, each link l_α is endowed with an external factor λ_{l_α} randomly extracted from the same exponential distribution as λ_α . The dynamics starts by connecting the

²⁶ For example, when there exist specialized interactions among two nodes with degree one. Such a link will never be rewired (as it would leave one of the two nodes with degree zero). This pathological case may appear either during the algorithmic process, or in the initial configuration when one starts from an empirical network.

first plant α_1 with k_{α_1} animals. For the other plants α and their links l_α , if $r_\alpha > \lambda_{l_\alpha}$, then the plant connects with the first animal in \mathcal{A}' , the set of animals that have not yet been linked to by another plant $\beta \neq \alpha$. Conversely, if $r_\alpha \leq \lambda_{l_\alpha}$, then the plant connects with an animal in \mathcal{A}'' , the set of animals that already gained at least one connection in previous time steps. If the plant has already connected to all the animals in \mathcal{A}' or \mathcal{A}'' , it only connects to animals in the non-saturated set.

The model can reproduce more than 70% of the structural metrics (degree distribution, modularity, and nestedness) of empirical networks [27]. This percentage is much larger than that achieved by previous models based on trait-matching [256] and preferential attachment [307]. For both pollination networks and manufacturer-contractor networks from the New York garment industry, the model can generate networks with levels of nestedness compatible with the empirically observed ones (see Fig. 2 in [27]).

5.2.3. Fitness maximization based on population dynamics

In an adaptive system, it is plausible to consider a mechanism where a node tends to delete its less profitable connection and to replace it with a new, more profitable connection. We have already leveraged this idea when introducing the Self-organizing Network Model (SNM, Section 5.2.1), and we shall explore analogous mechanisms in the context of socio-economic systems (see Section 5.3.3). When assuming that nodes strive to maximize their centrality, or fitness, the resulting model critically depends on how we define node centrality or fitness in the first place. In this Section, we focus on fitness-maximization mechanisms where species strive to maximize their abundance as determined by the stationary state of a suitable population dynamics model [33,299]. This class of mechanisms is akin to variational principles in physics, where one seeks to find the optimal state for a given system by maximizing (or minimizing) a suitable function of the variables that describe the system's configuration. As an example, the equilibrium state of a thermodynamic system can be found by maximizing the entropy of the probability distribution of its configurations [168].

Zhang et al. [299] proposed a rewiring mechanism where, starting from a given topology, at each step, a species is randomly selected. The selected species deletes its connection that brings the smallest increase in its per-capita population, as determined by the stationary state of a model of mutualistic population dynamics (see [299] for the details). The deleted connection is replaced ("interaction switch" [299]) by a randomly chosen connection. Zhang et al. [299] found that this simple behavioral mechanism leads to the emergence of nested structures that exhibit levels of nestedness relatively close to those observed in real networks.

Suweis et al. [33] studied a dynamics where, starting from a random species-species interaction matrix, the species rewire their interactions with the goal to maximize their abundance. They showed that such dynamics results in a nested structure of the interaction network. Fitness optimization, therefore, may be a suitable explanation of the nested patterns found in mutualistic networks. In the following, we provide the essential elements of the work by Suweis et al. [33], referring to the original paper for all the details and additional insights.

Linear and non-linear population dynamics. To evaluate the impact of an interaction swap on the individual species' abundances and on the community's total population, Suweis et al. [33] considered both a linear dynamics and a non-linear population dynamics. We denote by N and M the number of animal and plant species, respectively; we denote by $\{x_1, \dots, x_N\}$ and $\{x_{N+1}, \dots, x_{N+M}\}$ the abundances of the animal and plant species, respectively.²⁷ One can define an interaction matrix \mathbf{M} composed of four blocks:

$$\mathbf{M} = \begin{bmatrix} \Omega^{AA} & \Gamma^{AP} \\ \Gamma^{PA} & \Omega^{PP} \end{bmatrix} \quad (66)$$

The blocks Ω^{PP} and Ω^{AA} represent the plant-plant and animal-animal competitive interactions, respectively; the blocks Γ^{PA} and Γ^{AP} represent the plant-animal mutualistic interactions. The linear dynamics is given by [33]

$$\frac{d\mathbf{x}}{dt} = \mathbf{x}(\boldsymbol{\alpha} - \mathbf{M}\mathbf{x}), \quad (67)$$

where $\boldsymbol{\alpha}$ represents the vector of intrinsic growth rates. The non-linear dynamics is described by the equation [33]

$$\frac{d\mathbf{x}}{dt} = \mathbf{x} \left(\boldsymbol{\alpha} - d\mathbf{x} - \frac{\mathbf{M}\mathbf{x}}{h + \Theta\mathbf{x}} \right), \quad (68)$$

where d represents a self-interaction strength, h is a parameter referred to as *handling time*,²⁸ and Θ is a matrix whose elements are defined as $\Theta_{ij} = 1$ if i and j are neither both plants nor both animals, $\Theta_{ij} = 0$ otherwise. It is worth noticing that this model can be obtained from the model with competition and mutualism studied in [38,50] (see Eq. (89) and the related discussion below) by switching off interspecific competitive interactions. One can write down the equations for the stationary point and the local stability conditions for this dynamics — we refer to the Supplementary Information of [33] for details.

²⁷ This notation is different from the one that we defined in Table 2 where we labeled animals and plants through Latin ($i \in \{1, \dots, N\}$) and Greek letters ($\alpha \in \{1, \dots, M\}$), respectively. However, we adopt it in this Section as it is convenient for a concise description of the model.

²⁸ The handling time is defined as the amount of time needed for the plant-animal interaction to take place, which limits the total benefits for a given species from the interactions with its mutualistic partners. We refer to Section 6.2.3 for more details about this mutualistic model and its implications for systemic stability and feasibility.

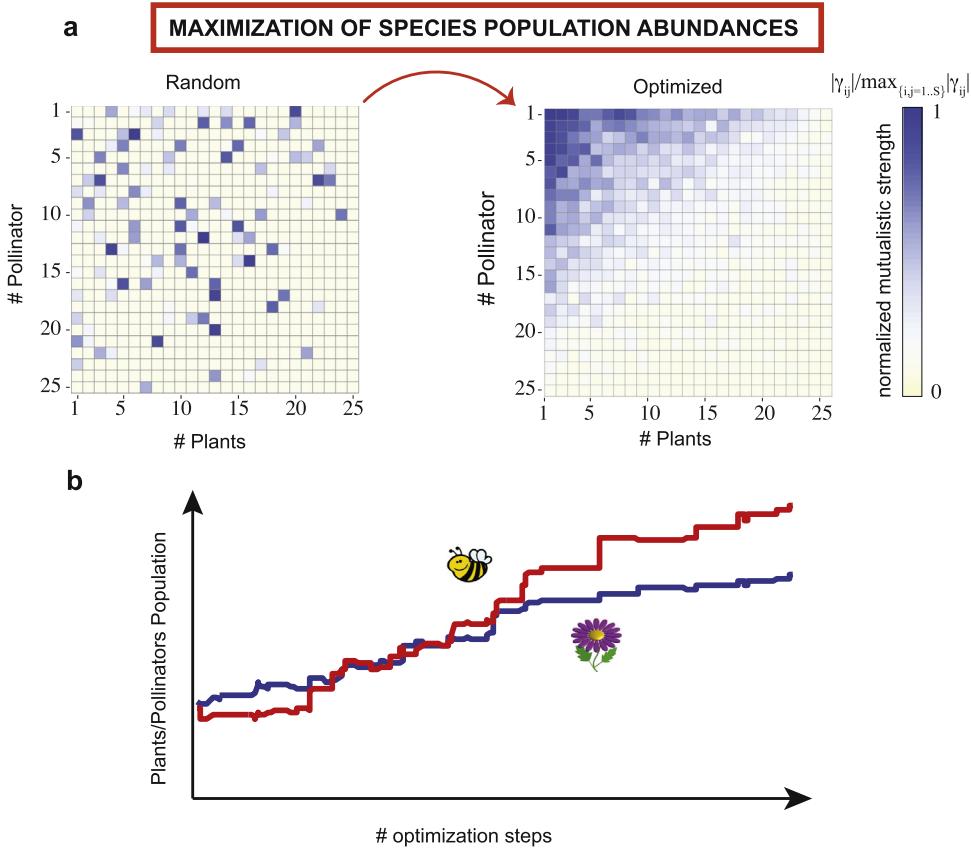


Fig. 20. Effects of the species-abundance maximization dynamics by Suweis et al. [33]. As the number of optimization steps increases, **a** the structure of the interaction matrix gradually approaches a nested structure; **b** the population of both plants and pollinators increases.
Source: Reprinted from [33].

Community-level and species-level optimization dynamics. The *community-level optimization* dynamics starts with $x_i = 1$ for each species i , and a randomly generated interaction matrix M . We assume that the vector α of intrinsic growth rates stays constant during the optimization dynamics. At each time step n of the optimization dynamics, we randomly select a pair (i, j) of interacting species ($\gamma_{ij} \neq 0$) and a pair (l, m) of non-interacting species ($\gamma_{lm} = 0, (l, m) \neq (i, j)$). We propose a swap between the two interactions: $\tilde{\gamma}_{lm}^{(n+1)} = \gamma_{ij}^{(n)}$ and $\tilde{\gamma}_{ij}^{(n+1)} = \gamma_{lm}^{(n)}$. To decide whether to accept the swap, we compute the new equilibrium point of the system: in the linear dynamics, $\tilde{x}^*(n+1) = [\tilde{M}(n+1)]^{-1} \alpha$, where the two original matrix elements have been swapped in \tilde{M} . We accept the swap only if $\sum_{i=1}^{A+P} \tilde{x}_i^{*(n+1)} \geq \sum_{i=1}^{A+P} x_i^*(n)$; in this case, $\Gamma(n+1) = \tilde{\Gamma}(n+1)$. If, by contrast, $\sum_{i=1}^{A+P} \tilde{x}_i^{*(n+1)} < \sum_{i=1}^{A+P} x_i^*(n)$, we set $\Gamma(n+1) = \Gamma(n)$ (see Fig. 20).

This community-level optimization dynamics corresponds to a rewiring of the interaction matrix Γ aimed at maximizing the total population $\sum_i x_i$ of the system. It can be interpreted as a group selection process, as opposed to a process where species aim to maximize their own abundance. The latter mechanism can be included by accepting a potential interaction swap not if the system's total population increases, but if the abundance of the species involved in the interaction increases (*species-level optimization*). More specifically, at each step n of the optimization process, one selects a species i and two of its links $(i, j), (i, k)$. Again, one proposes an interaction switch $\tilde{\gamma}_{ij}^{(n+1)} = \gamma_{ik}^{(n)}$ and $\tilde{\gamma}_{ik}^{(n+1)} = \gamma_{ij}^{(n)}$. As opposed to the community-level optimization process, the swap is only accepted if it would not result in a decrease of species i 's abundance (as determined by the steady state of the population dynamics under consideration).

Importantly, both optimization dynamics drive the interaction matrix from the initial random topology to a highly nested topology. To analytically understand this phenomenon, one can consider a mean-field scenario where interaction and competition strength parameters are homogeneous across all pairs of species. Within the mean-field scenario, the fact that the above-described optimization processes lead to nested topologies is a consequence of two properties: (1) both optimization dynamics presented above tend to increase the total population of the system²⁹; (2) the total population of

²⁹ While this is true by construction for the community-level optimization, it may look counterintuitive for the species-level optimization process. If species aim at maximizing their own abundance, why does the total population increase? To answer this question, one can use a perturbative

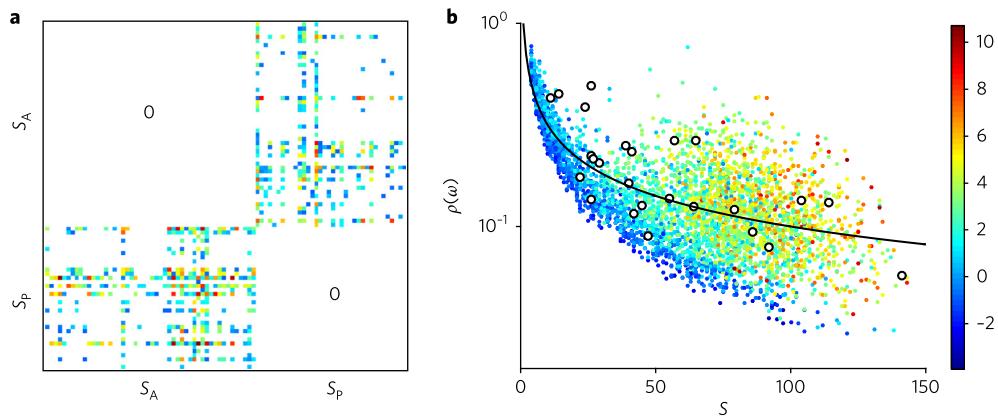


Fig. 21. Nestedness in the speciation–divergence model. Panel **a** illustrates a bipartite adjacency matrix obtained with the model. Panel **b** represents the spectral radius $\rho(\mathbf{A})$ (see Section 3.1.4) of 4000 networks generated with the speciation–divergence model (for different values of the input parameters) as a function of the number of species S . The circled dots correspond to real datasets, whereas the continuous line represents the analytic expectation for a random network.

Source: Reprinted from [34].

the system is significantly correlated with the nestedness of the interaction matrix. In the mean-field scenario, the total population x^{tot} of the system ($x^{tot} = \sum_i x_i$) depends linearly on nestedness according to a relation $x^{tot} \simeq B_1 O^{tot} + B_2$, where O^{tot} is the total overlap of the interaction matrix (proportional to the sum of all the node–node pairwise overlaps $O_{ij} = \sum_l A_{il} A_{jl}$), and B_1, B_2 are two constants that depend on the interaction parameters γ, ω , on the network connectance C and on the number of species S . As nestedness (as measured by NODF) increases with the total overlap O^{tot} of the interaction matrix, it follows that nestedness increases with the total population of the system (see Fig. 20 and [33]).

5.2.4. Nestedness as an evolutionary spandrel: Speciation and divergence

Valverde et al. [34] recently introduced a simple generative model for mutualistic networks based on two mechanisms: speciation and divergence. The model focuses on weighted networks, where the weight of an interaction between animals and plants may represent the number of observed interaction within a given time interval. In an evolutionary ecology spirit, the *speciation* process is essentially a duplication process where a species splits into two separate species. In the model, with a given probability π_A , an animal species i undergoes a *speciation* process where a new (daughter) animal species j is created with the same links of the original (parent) species i : if we denote as $\omega_{i\alpha}$ the weight of the interaction between animal i and plant α , we have $\omega_{i\alpha} = \omega_{j\alpha}$ for all plant species $\alpha \in \{1, \dots, M\}$. With a given probability π_B , a plant species β undergoes an analogous speciation process.

Species *divergence* is basically a weight randomization process. The weights between parent and daughter species are then redistributed. In particular, we generate a random number $\mu \in (0, 1)$ and we update the weights as follows: $\omega_{i\alpha} \rightarrow (1 - \mu) \omega_{i\alpha}, \omega_{j\alpha} \rightarrow \mu \omega_{j\alpha}$. Besides, with a given probability p , each link (i, α) can change weight: $\omega_{i\alpha} \rightarrow \omega_{i\alpha} + \xi$, where $\xi \in (\beta, \beta)$ is a relatively small random number, and p, β are parameters of the model. The model also features a maximum total weight for plants and animals, and a threshold minimal value for links' weights.

The networks generated according to this rule exhibit various degrees of nestedness, from low to high, conditional to the values of the input parameters (Fig. 21). Based on this result, Valverde et al. [34] pointed out that network nestedness and other structural properties (such as the network's heterogeneous degree and weight distribution, and link asymmetry [34]) can be viewed as “evolutionary spandrels”, i.e., as structures that are the consequence of the system's building rules, have well-defined non-random features, and reveal some of the underlying rules of construction. Importantly, these findings indicate that nestedness can emerge as a result of building rules that do not involve species-level or community-level optimization, as opposed to the optimization rules that we investigated in the previous Sections.

Evolutionary mechanisms to explain the emergence of nestedness were also previously considered by Takemoto and Arita [308]. They assumed that starting from a small initial network, new plants can emerge from randomly selected existing ones through mutation. The traits of the new plants are similar to those of the original ones. Based on the assumption that plant–animal interactions are driven by trait-matching mechanisms, some of the animals that interact with the new plant are inherited from those that interacted with the ancestral plant. Besides, a plant can acquire new traits

argument to show that in the limit $\gamma \sim \omega \ll 1$, at each optimization step, the total population increase is approximately equal to the abundance increase of the species involved in the rewiring. Increasing the abundance of the species involved in the interaction, therefore, results in increasing the total population of the system. We refer the reader interested in this intriguing equivalence to Section 4.3 and 5 of the Supplementary Information of [33].

and new, randomly selected partners. This mechanism generates nested networks as well, and the resulting networks were reported [308] to match the nestedness of empirical networks more accurately than those generated by the bipartite cooperation model described in Section 5.2.2.

5.2.5. Invasion dynamics with different assemblage rules

The results by Valverde et al. [34] indicate that network formation mechanisms without optimization can explain the emergence of nestedness in mutualistic networks. In a similar spirit, Maynard et al. [35] considered a community of N competing species. They were interested in finding slightly different stochastic mechanisms of network growth that lead to drastically different network topologies. If such similar but different mechanisms can be identified, one might argue that widely different observed topologies might be the result of different assemblage rules, without the need for introducing formation mechanisms based on optimization or selection for stability.

By denoting as $x_i = x_i(t)$ the abundance of species i at time t , they considered a competitive dynamics of the form

$$\frac{dx_i}{dt} = x_i(t) \left(r_i + \alpha_i \sum_j A_{ij} x_j \right), \quad (69)$$

where r_i represents the intrinsic growth rate and $\alpha_i < 0$ represents the competition parameter. They considered a simple network formation process where one starts with a single species and, at each step, considers a potential “invader”. They introduced two mechanisms, referred to as *immigration* and *radiation*, that differ in the way the invader interacts with the preexisting species (see next paragraph). If the potential invader satisfies the invasion condition (see Eq. (70) below), it is added to the system. Once it has been added to the system, one runs the dynamics described by Eq. (69) until a stable point is reached. The addition of the invader perturbs the system, and some species might become extinct as a result. Once the stationary point is reached, one repeats this procedure by proposing a new invader.

An invader i is added to the system if and only if

$$r_i + \alpha_i \sum_j A_{ij} x_j > 0. \quad (70)$$

The validity of this condition depends on the coefficients A_{ij} . Maynard et al. [35] considered two mechanisms:

- In the *immigration* scenario, the A_{ij} coefficients for the potential invader i are random numbers extracted from the uniform distribution $\mathcal{U}(-1, 0)$ for $j \neq i$, whereas $A_{ii} = 1$. If the potential invader satisfies the invasion condition (Eq. (70)), then it is added to the system. Otherwise, one replaces randomly-chosen coefficients A_{ij} with new random numbers extracted from $\mathcal{U}(-1, 0)$ until the invasion condition is met and the invader can enter the system.
- The *radiation* scenario differs from the immigration scenario in the way the coefficients are replaced if the invasion condition is not met. To allow the potential invader i to enter the system, in the radiation scenario, one randomly selects preexisting species and slightly perturbs their coefficients, until the invasion condition for species i is met.

The two mechanisms lead to drastically different levels of nestedness in the generated interaction networks. While the nestedness of immigration-generated networks is not distinguishable from that of randomized networks, the nestedness of radiation-generated networks tends to be significantly larger than that of their randomized counterparts. This points out that: (1) different assemblage mechanisms can lead to different network topologies, without the need for considering optimization mechanisms; (2) Radiation-like assemblage mechanisms might be responsible for the levels of nestedness observed in empirical networks. We will come back to these points in Section 5.4.

5.3. Social and economic mechanisms

Compared to ecological networks, scholars have started investigating the emergence of nestedness in socio-economic networks much more recently. In this Section, we focus on network formation and rewiring mechanisms that aim to explain the emergence of nestedness in unipartite socio-economic networks (Sections 5.3.1–5.3.3) and in bipartite country–product networks (Section 5.3.4).

5.3.1. Network formation based on centrality maximization

In [36], the authors introduced a network formation model where the agents attempt to maximize strategically their centrality in the network in presence of link decay. Therefore, it is a network formation model where the number of edges changes over time until it reaches a stationary value which depends on the model parameter α (see below). In that dynamics, one starts from an arbitrary topology. At each time step, a randomly selected node has two options. With probability α , it creates a link to the most central node that is not already among its neighbors. With probability $1 - \alpha$, it deletes its connection with the least-central of its neighbors. The dynamics has one parameter, α , and it converges to a stationary, perfectly nested topology, provided that the system started from a perfectly nested topology.

The authors in [36] noted that this evolutionary dynamics leads to a nested network independently of the exact notion of centrality that is used. Later, Schoch et al. [309] showed formally that in a nested network all notions of centrality are ranked in the same way in a nested network. An insightful consequence is that agents do not become central because of

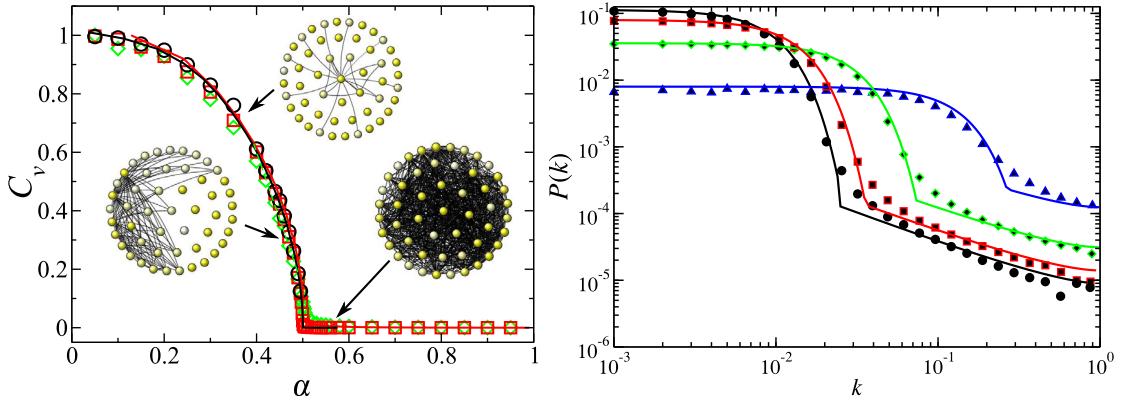


Fig. 22. Transition from a fully connected topology ($\alpha > 1/2$) to a nested, highly centralized topology ($\alpha < 1/2$) in the growing network model based on centrality maximization [36]. The transition manifests itself in the discontinuity of the derivative of the eigenvector centralization C_v at the critical point $\alpha_c = 1/2$ (panel (a)). The degree distributions of the networks generated by the model (panel (b)) exhibit a two-step relaxation which combines an exponential (low k) and a power-law (large k); the different curves in panel (b) correspond to different values of α .
Source: Reprinted from [36].

their actions; they do so because they turn into the best prospective connection for other agents who seek to increase their own centrality. This simple mechanism reinforces the nestedness in the network; concurrently, it promotes central agents to become even more central, and it confines peripheral agents to remain in those positions. Further, when the network is nested, agents can find the best connection target even with local information.

In the model, one can write down the master equation [310] for the degree distribution $p_k(t)$ for the nodes in the independent sets as³⁰ [36]

$$\frac{\partial p_k(t)}{\partial t} = \mathbb{P}[k+1 \rightarrow k] p_{k+1}(t) + \mathbb{P}[k-1 \rightarrow k] p_{k-1}(t) - (\mathbb{P}[k \rightarrow k-1] + \mathbb{P}[k \rightarrow k+1]) p_k(t), \quad (71)$$

where $\mathbb{P}[k+1 \rightarrow k] = (1-\alpha)/N$ is the probability that a node deletes one connection in the time unit, whereas $\mathbb{P}[k-1 \rightarrow k] = \alpha/N$ is the probability that a node acquires one connection. The smallest degree increase is given by $\delta k = 1/N$. One can therefore expand the previous Equation in powers of δk , obtaining (for $k > 0$) [36]

$$\frac{\partial p_k(t)}{\partial t} = (1-2\alpha) \frac{\partial p_k(t)}{\partial k} + \delta k \frac{\partial^2 p_k(t)}{\partial k^2} + \mathcal{O}(\delta k^2) \quad (72)$$

For $\delta k \ll |1-2\alpha|$, the $\mathcal{O}(\delta k)$ diffusion term can be neglected and we are left with a drift equation whose asymptotic solution is either a fully-connected network ($\alpha > 1/2$) or an empty network ($\alpha < 1/2$). Around $|1-2\alpha| \simeq \delta k$, i.e., when $\alpha \simeq 1/2 + \mathcal{O}(1/N)$, the diffusion term is not anymore negligible, and the dynamics is described by the Fokker–Planck equation

$$\frac{\partial p_k(t)}{\partial t} = (1-2\alpha) \frac{\partial p_k(t)}{\partial k} + \frac{\partial^2 p_k(t)}{\partial k^2}. \quad (73)$$

In the regime $\alpha < 1/2$, one solves Eq. (72) to derive the full degree distribution p_k . One finds that p_k has a two-step behavior: the distribution is exponential up to a critical degree value k^* which can be analytically determined, whereas it exhibits a power-law tail $\sim k^{-1}$ for $k > k^*$ (see Fig. 22b).

It is interesting to look at the centralization of the eigenvector centrality vector as a function of the model parameter α (see Fig. 22a). Intriguingly, the model exhibits a sharp phase transition at $\alpha = 1/2$: the system has an even degree distribution for $\alpha > 1/2$, which results in a centralization of the eigenvector centrality vector close to zero. By contrast, the system becomes highly-centralized for $\alpha < 1/2$, which results in a significant level of nestedness [36]. To conclude, it is worth noticing the analogy between social-climbing models and ecological network formation models where species attempt to maximize their fitness (like the SNM, see Section 5.2.1), which points out that explaining the same structural pattern across different systems might unveil common formation mechanisms.

5.3.2. Network formation in a two-stage game

Later, in [25], a dynamic network formation model was developed that can explain the observed nestedness of real-world economic networks: e.g., interbank loans, trade in conventional goods and arms trade between countries (cf. Fig. 2). Within the framework of the random utility model [311], links in the network are formed on the basis of agents' centrality

³⁰ As in [36] we neglect all the contributions from the nodes that belong to the dominating set. These contributions are nevertheless $\mathcal{O}(1/N)$ small.

and have an exponentially distributed lifetime. In the model, at each period of time, agents play a two-stage game. In the first stage, as in [312], agents compute their effort level which is proportional to their *Bonacich centrality* [228]. In the second stage, a randomly chosen agent can update her linking strategy by creating a new link as the best response to the current network. Links do not last forever but have an exponentially distributed life time. The most valuable links (i.e. the ones with the highest Bonacich centrality) decay at a slower rate than those that are less valuable. As a result, this model considers the formation of economic networks as the result of a tension between search for new linking opportunities and volatility that leads to the decay of existing links.

The dynamics runs as follows. The system is composed of N agents. At any time t , there is a network $\mathcal{G}(t)$ with adjacency matrix \mathbf{A} that describes the connections between agents. Then, the two-stages of the game are applied. In the first stage, each player selects an effort level x_i to put into the game. It is assumed that the payoff π_i she receives is

$$\pi_i = x_i - \frac{1}{2}x_i^2 + \lambda \sum_{j=1}^N A_{ij}x_jx_j, \quad (74)$$

where λ is a parameter that weights the relative contribution of the idiosyncratic component and network interdependencies to the total payoff. Ballester et al. [312] found that the equilibrium effort levels that maximize the payoff of participant i is a function of i 's Bonacich centrality $b_i(\lambda)$: $\pi_i^*(\mathcal{G}, \lambda) = b_i(\lambda)/2$. In this stage of the game, agents select their equilibrium effort levels.

In the second stage, the network topology is updated. First, an agent i has the opportunity to create a link (i, j) – i.e., the network $\mathcal{G}(t)$ transitions to another one $\mathcal{G}(t) \oplus (i, j)$ – with a rate

$$v_+(\mathcal{G}(t) \oplus (i, j)) \propto \exp(\pi_i^*(\mathcal{G}(t) \oplus (i, j), \lambda)/\zeta). \quad (75)$$

Second, links have an expected lifetime which is exponentially distributed. In particular, the rate of link decay becomes

$$v_-(\mathcal{G}(t) \ominus (i, j)) \propto \exp(\pi_i^*(\mathcal{G}(t) \ominus (i, j), \lambda)/\zeta). \quad (76)$$

The parameter ζ controls the level of randomness in the network dynamics – essentially, it plays the role of an inverse temperature.

Stochastic stability was used to identify the network's stationary state; the authors [25] found that the resulting network topology is fully nested when no randomness exists in the agent's decisions. The degree of nestedness is preserved to a large extent for small temperature. In particular, in the limit $\zeta \rightarrow \infty$, we recover the centrality-maximization dynamics [36] described in Section 5.3.1.

5.3.3. Network rewiring based on social climbing

In contrast to the previous models, the social climbing game [81] considers a static system composed of a fixed number of nodes and links. This model is a highly stylized model of a society where individuals seek to maximize their social status by forming new connections and deleting old ones in a strive to maximize the centrality of their interaction partners. The essential role of centrality as a driver for success has been proven in several studies from different angles. For instance, central individuals in the social network tend to score higher in education tests [313]; individuals' centrality in communication networks is a predictor for employees' future promotions and resignations [314–316]; central individuals tend to have the largest impact on the opinions of their peers according to linear opinion formation models [317]. In light of these studies, the assumption that individuals are, to some extent, driven by the maximization of their centrality is a reasonable one [81,318].

Consider a system composed of N individuals and a fixed number E of undirected links between them. Bardoscia et al. [81] introduced a model where each agent i is characterized by a “social capital” utility function, u_i , defined as

$$u_i = \sum_j A_{ij} k_j + \mu k_i, \quad (77)$$

where μ is a parameter of the model that rules the relative strength of the two terms. At each step of the dynamics, one selects randomly a node i together with one of its neighbors (i.e., a node j such that $A_{ij} = 1$) and one of the selected neighbor's neighbors (a node l such that $A_{lj} = 1$). If i was already connected to l , nothing happens. If i was not connected to l , the link (i, j) is replaced by (i, l) with probability

$$\mathbb{P}[(i, j) \rightarrow (i, l)] = \frac{e^{\beta \Delta u_i}}{1 + e^{\beta \Delta u_i}}, \quad (78)$$

where Δu_i denotes the variation in node i 's utility caused by the possible link replacement.

The rewiring probability defined by Eq. (78) encodes the tendency of individuals to connect to individuals who bring them large utility gains. The strength of this tendency is ruled by the intensity parameter β , which can be interpreted as an inverse temperature, as we shall see below. In a “frozen” society ($\beta \gg 1$), the nodes always rewire their connections when this leads to an increase of their utility. In a “high-temperature” society ($\beta \simeq 0$), the rewiring probability approaches 1/2 regardless of the utility gain Δu , which makes the rewiring of links essentially independent of the candidate new neighbor's centrality.

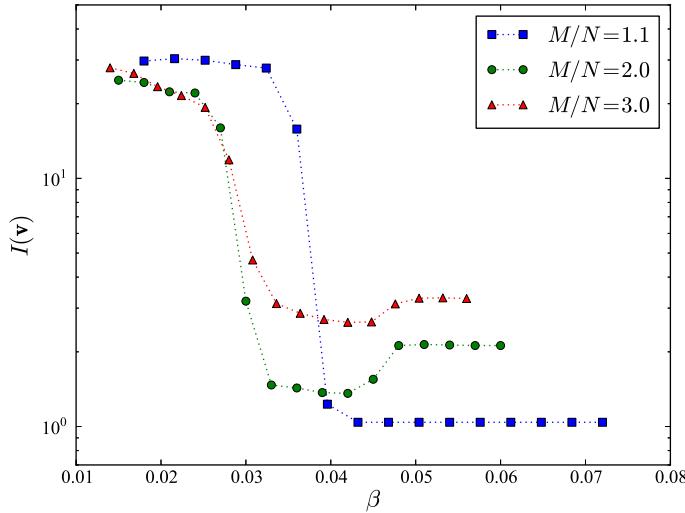


Fig. 23. Transition from an even, “disordered” topology (high-temperature, low β) to a highly centralized topology (low-temperature, high β) in the social climbing rewiring dynamics [81]. The transition manifests itself in the drop of the inverse participation ratio I of the network degree sequence as a function of the inverse temperature β ; different symbols stand for different values of average degree $\langle k \rangle = 2L/N$.
Source: Reprinted from [81].

The interpretation of the intensity parameter β as the inverse temperature of a statistical-mechanics system is a compelling feature of the model. To understand it, one can show that (1) when a rewiring $(i, j) \rightarrow (i, l)$ is performed, the variation ΔU of the total utility $U := \sum_m u_m$ can be expressed as twice the variation Δu_i of node i 's local utility: $\Delta U = 2 \Delta u_i$ [81]; (2) the dynamics described above is ergodic. Hence, the dynamics based on Eq. (78) converges to the equilibrium of a system described by the Hamiltonian function [81]

$$H = -U = -\sum_i k_i^2 - \mu \sum_i k_i = -N \langle k^2 \rangle - N \mu \langle k \rangle, \quad (79)$$

with fixed average degree $\langle k \rangle$ and temperature $2/\beta$. In contrast to the dynamic processes described in Sections 5.3.1 and 5.3.2, this one preserves the network size and density throughout the dynamics.

The results of numerical simulations based on this model show that as β increases, the system abruptly transitions from a rather even degree distribution to a highly unequal degree distribution where few individuals are connected with the majority of the others, whereas most of the individuals have only few connections. A useful quantity to visualize this transition is the degree-sequence's inverse participation ratio $I(\mathbf{v})$ – i.e., a proxy for the effective number of nodes who received connections³¹ – which displays an abrupt jump from a $\mathcal{O}(N)$ value to a $\mathcal{O}(1)$ value (see Fig. 23). Besides, high- β (low-temperature) systems are also characterized by an increased temporal stability of the ranking of the nodes by degree (see Fig. 6 in [81]), which is representative of a low-mobility society where pre-defined hierarchies are likely to be preserved. We have seen that “low-temperature” nested structures can be interpreted as the fingerprints of “order” in biogeographic systems [22], see Section 3.1.2; similarly, highly nested hierarchies in society can be interpreted as the manifestation of a “low-temperature” societal dynamics where it is hard for an individual to climb well-established ladders, regardless of his/her individual capabilities.

5.3.4. Emergence of nestedness in the country–product network: Innovation and novelty

While capability-based models [29,52] (see Section 4.3.4) can generate nested country–product export networks, they do not describe the detailed dynamics that leads the countries to diversify their export baskets. To describe this process, in a similar spirit to non-linear preferential attachment models in network science [319], Saracco et al. [51] introduced a network formation model with three steps:

- A country is selected with probability $P_i \sim k_i^a$, where $a > 0$ is a parameter of the model. Such probability reflects the intuition that more diversified countries are more likely to further diversify their export basket.
- The chosen country i selects a product α that already belongs to its export basket according to the probability $P(\alpha|i) \sim k_\alpha^b$, where $b > 0$ is a parameter of the model. The rationale behind this rule is that the more ubiquitous a

³¹ Given a normalized N -component vector \mathbf{v} ($\sum_i v_i^2 = 1$), its inverse participation ratio $I(\mathbf{v})$ is given by $I(\mathbf{v}) = (\sum_i v_i^4)^{-1}$. A vector that is completely localized (i.e., a vector that has one component equal to one and all the others equal to zero) has $I(\mathbf{v}) = 1$. A fully delocalized vector (i.e., a vector whose components are all equal to $1/\sqrt{N}$) has $I(\mathbf{v}) = N$.

product is, the more countries possess the apposite capabilities to make it, the higher the possibility of generating an innovation from that product.

- Given the selected country–product pair (i, α) , the country has now the last choice to make. It considers two types of potential new products: all the products $\tilde{\alpha}$ that belong to the neighborhood of α in the product space [286,287], and that were not already exported by country i ; a new product α'_{new} sprouting out from α . Within this set $\{\alpha', \alpha'_{new}\}$ of candidate products, country i chooses the new product $\tilde{\alpha}$ to export according to the probability $P(\tilde{\alpha}|i, \alpha) \sim (k_{\tilde{\alpha}} + k_0)^c$, where $c > 0$ and $k_0 > 0$ are parameters of the model.

The model has four parameters: a, b, c, k_0 . The parameter k_0 determines whether countries tend to move in the existing adjacent of already exported products (diversification, small values of k_0), or rather prefer to innovate and introduce new products in the system (innovation, large k_0). The model can generate bipartite country–product networks whose values of row-nestedness and column-nestedness (see Section 3.1.3) are compatible with those observed in the real networks (see Fig. 4 in [51]).

5.4. Bottom-line: Why does nestedness emerge in empirical networks?

As we emphasized in Section 5.2, the fact alone that a topology can theoretically arise as a result of a given dynamics does not guarantee that the studied dynamics shaped the emergence of that topology in real systems. The mechanisms investigated above are all able to generate nested networks, yet it remains unclear which of them best describe the formation of real-world systems. In both ecological and socio-economic networks, this uncertainty can be ascribed to the focus of most extant studies on static, time-aggregate datasets of ecological communities and trade networks, and on processes that can theoretically lead to the observed topologies. In ecology, recent works [320,321] have started investigating the temporal dynamics of communities. Crucially, increasing efforts in this direction may allow us to single out those mechanisms that likely led to the emergence of the observed topologies, while ruling out mechanisms that lead to realistic network topologies but do not reproduce the observed network dynamics.

From an interdisciplinary standpoint, we envision that the validation of network formation mechanisms (and, therefore, of potential explanations for nestedness) will likely be performed soon in socio-economic networks, where typically one has access to fine-grained temporal data. Existing studies on socio-economic networks have both introduced candidate mechanisms to explain the emergence of nestedness (Section 5.3), and studied the detailed temporal dynamics of nestedness (see Section 4.1.4 and, in particular, Fig. 16). Validating the proposed mechanisms through robust statistical analysis, perhaps in a similar way to procedures adopted for growing information networks [255], is a natural step for future research.

6. Implications of nestedness for systemic stability and feasibility

In occasion of the 100th anniversary of the British Ecological Society, Sutherland et al. [322] have collected a list of 100 fundamental ecological questions whose solutions have the highest potential to advance ecological science. Among them, we find: "How does the structure of ecological interaction networks affect ecosystem functioning and stability"? [322] The main goal of this Section is to address this question for nested structures: how does nestedness impact on the robustness of the system against extinctions and targeted attacks? Does nestedness facilitate or harm the stable co-existence of species?

Driven by these questions, we focus here on two aspects of systemic stability: network *robustness against external perturbations* and *dynamical stability*. Network robustness against structural perturbations (Section 6.1) is typically defined in terms of node-removal processes. The core idea behind it is that when we remove a node in the network, other nodes might become extinct because they were only connected with the removed node – a phenomenon known as "co-extinction" [323]. Sequentially removing all the nodes (either in order of decreasing degree or in order of decreasing score as determined by some ranking algorithm) therefore leads to a process where eventually all the nodes become extinct.

Besides, the topology of an interaction network has important consequences on dynamical processes that act on it. Ecologists are especially interested in revealing whether some widely-observed network structural patterns can increase the dynamical stability of the system, i.e., its capability to return at its equilibrium point after a perturbation of the species' population. They are also interested in whether observed topological patterns facilitate the co-existence of species and, therefore, the community's feasibility. The impact of nestedness on dynamical stability and feasibility has been studied both with a population-dynamics approach [38,50,324,325] and with a random-matrix theory approach [37,326].

In the population-dynamics approach, one considers a dynamic model with competition and mutualism (which acts on the underlying species–species interaction network), and seeks to find whether different network topologies facilitate or impede the stable co-existence of species. We will review this approach in Section 6.2. We will then focus on the results from random-matrix theory on the local stability of random, mutualistic, and nested interaction matrices (Section 6.3). We will finally review the obtained results within the various approaches and draw a bottom-line on the impact of nestedness on systemic stability and feasibility, based on the results obtained so far (Section 6.4).

6.1. Robustness against structural perturbations

Understanding the co-extinction patterns triggered by the loss of a given species is central to the preservation of complex systems. In ecological systems, it can indeed inform decisions on which species to give higher priority for research and management in order to preserve a given ecosystem — we refer to [327] for a review on the application of ecological networks to applied conservation practices. In economic and financial systems, understanding how the failure of an actor (like a financial institution) can trigger large-scale cascades of failures is essential to prevent the collapse of the global financial system [328,329].

Importantly, the topology of a network affects its robustness against structural perturbations, which can be performed by sequentially removing the nodes according to pre-established strategies. In the following, we investigate two main questions related to topological robustness: (1) When we remove sequentially the nodes in order of degree, do nested arrangement lead to a faster or a slower systemic breakdown? (2) Do more nested arrangements of the adjacency matrix (as produced by non-linear ranking algorithms introduced in Section 3.3.2) lead to faster or slower systemic breakdown? The two questions will be addressed in Sections 6.1.1 and 6.1.2, respectively.

6.1.1. Topological robustness of nested networks

Consider a bipartite network. We label the two kinds of nodes as “active” and “passive species”, yet the lines of reasoning presented in this section equally apply to any bipartite network. If we delete some active (passive) species from the network, some passive (active) will become extinct as a result. The loss of single species can even trigger the co-extinction of several other species. Such a process, called *co-extinction cascade*, has been extensively studied in the ecological literature [305,330]. In this Section, we focus on the impact of a nested structure on the speed of network disruption under such dynamics. In particular, scholars have debated whether an increase in the level of nestedness in a given network is associated with an enhanced robustness. We shall see in the following that this property only holds for specific removal strategies of the nodes.

Quantifying systemic robustness: attack-tolerance curve (ATC). To quantify the impact of network topology on network robustness under node deletion processes, we introduce the attack-tolerance curve (ATC) [330]. This curve represents the fraction of surviving passive species as a function of the fraction of deleted active species. The systemic robustness R of a given network is defined as the area under the ATC. By definition, $R \in (0, 1)$. A “robust” system is characterized by R values relatively close to one. This corresponds to a system where, even when a substantial number of active species is deleted, a considerable number of passive species are still alive. By contrast, a “fragile” system exhibits relatively small R values: even the removal of few active species is enough to substantially decrease the number of passive species in the system.

It follows from the definition of ATC that the robustness R depends not only on network topology, but also on the order in which the active species are sequentially deleted. Burgos et al. [305] considered three different removal strategies:

- $[+ \rightarrow -]$ (from generalists to specialists) removal strategy. The active species are removed in order of decreasing degree.
- $[- \rightarrow +]$ (from specialists to generalists) removal strategy. The active species are removed in order of increasing degree.
- Random removal strategy. The active species are removed in a random order.

While all the three cases are important from a theoretical standpoint, experimental evidence strongly suggests that in mutualistic networks, the specialist species have also higher likelihood to extinguish [331], which suggests that the $[- \rightarrow +]$ strategy is the most relevant one for this class of networks. Importantly, we shall see in the following paragraph that the $[- \rightarrow +]$ strategy is also the one for which the topological robustness benefits from a nested topology to the largest extent.³²

Nestedness and topological robustness. Memmott et al. [330] studied the co-extinction of species in plant–pollinator networks, and they found that the $[- \rightarrow +]$ removal strategy (from specialists to generalists) leads to a significantly slower network breakdown with respect to the random removal strategy and the $[+ \rightarrow -]$ (from generalists to specialists) removal strategy. They concluded that mutualistic networks are robust with respect to preferential node deletion, in particular when specialists have a higher probability to become extinct (see Fig. 24).

Burgos et al. [305] moved one step forward by using the self-organizing network model (SNM, see Section 5.2.1 and [80]) to gauge the impact of nestedness on systemic robustness. As we have seen in Section 5.2.1, the SNM allows us to gradually move from a completely random topology to a perfectly nested topology. The authors found that more nested structures are both more robust against the $[- \rightarrow +]$ removal strategy, and more fragile against the $[+ \rightarrow -]$ removal strategy (see Fig. 24). This is intuitive: in a large perfectly nested structure, if we lose the most specialist species α , no other species will become extinct as α was only connected with the most generalist species. By contrast, if we lose

³² Motivated by this consideration, Burgos et al. [332] introduced a nestedness metric based on the difference between the robustness coefficients corresponding to attack strategies $[- \rightarrow +]$ and $[+ \rightarrow -]$.

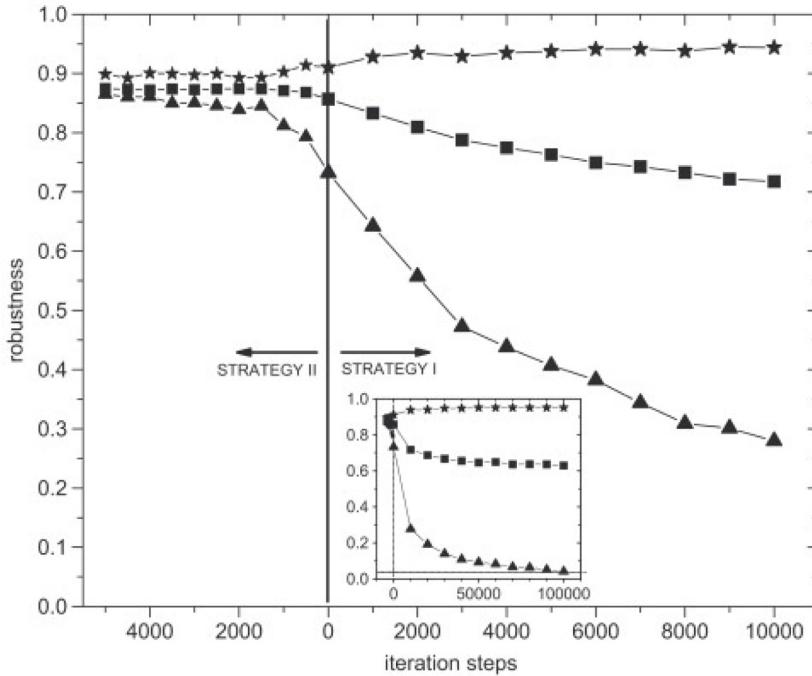


Fig. 24. Robustness as a function of the number of iterations of two rewiring strategies, for the mutualistic system studied by Clements and Long [333]. Strategy I corresponds to the Self-organizing Network Model (SNM, see Section 5.2.1); the nodes replace their partners with more generalist partners. This rewiring strategy gradually increases the level of nestedness in the system. Strategy II is the opposite: the nodes strive to connect with more specialist nodes. Increasing nestedness increases the robustness against $[- \rightarrow +]$ attacks (from specialists to generalists, filled stars), whereas it decreases the robustness against $[+ \rightarrow -]$ (filled triangles) and random attacks (filled squares). The inset focuses on Strategy I, by displaying results for a larger number of iterations.

Source: Reprinted from [305].

the most generalist species, the most specialist species are likely to become extinct as they were only connected with the most generalist species. Burgos et al. [305] noticed that if we accept that the only systems that we can observe in nature are those with a higher level of robustness, the omnipresence of nested structure in mutualistic networks might be a manifestation of the more vulnerable position of specialist species.

Incorporating interaction strength and link rewiring. While the removal strategies studied by Burgos et al. [305] revealed the implications of nestedness for the topological robustness of a given system, they are simplified descriptions of real species losses. Several subsequent studies have aimed at incorporating more realistic effects into the co-extinction cascade process. Kaiser-Bunbury et al. [334] pointed out that analyses like those carried out by Memmott et al. [330] and Burgos et al. [305] assume that all the removed species contribute equally to the fraction of removed species, regardless of their abundance and interaction frequency. To overcome this potential limitation, they defined the *normalized interaction strength* $L_{i\alpha}^{(N)}$ between a given animal i and a given plant α as the total number of visits per flower per hour, and the *absolute interaction strength* $L_{i\alpha}^{(A)}$ as $L_{i\alpha}^{(N)}$ times the mean number of flowers per square meters. The total interaction strength of animal i and plant α are defined as $N_i^{(A)} = \sum_{\alpha} L_{i\alpha}^{(A)}$ and $N_{\alpha}^{(A)} = \sum_i L_{i\alpha}^{(A)}$, respectively. The total interaction strength of the network is given by $N = \sum_i N_i^{(A)}$; in the co-extinction cascade process, the removal of species i leads to a loss $N_i^{(A)}$ in the total interaction strength.

Building on this setting, Kaiser-Bunbury et al. [334] analyzed 12 two-weeks snapshots of two pollination networks. They found that when the active species are removed progressively according to the strategy $[+ \rightarrow -]$, the attack tolerance curve of the total interaction strength $N^{(A)}$ as a function of the active-species total interaction strength displays a sigmoidal shape where the removal of few animals caused a sudden drop in the total interaction strength. Even more, they studied removal strategies that include the possibility of rewiring – the links observed across the 12 snapshot are considered as “potential links”, and these potential links were used as rewiring options for disrupted links during the removal process. Networks with potential rewiring turned out to be significantly more stable than non-rewired networks, which suggests that the possibility of rewiring might confer additional robustness to ecological communities.

6.1.2. Nestedness maximization and the structural importance of individual nodes

Section 6.1 dealt with the topological robustness of different network structures, for different node removal strategies. The analysis was restricted to removal strategies based on the ranking of the nodes by their degree. One can nevertheless

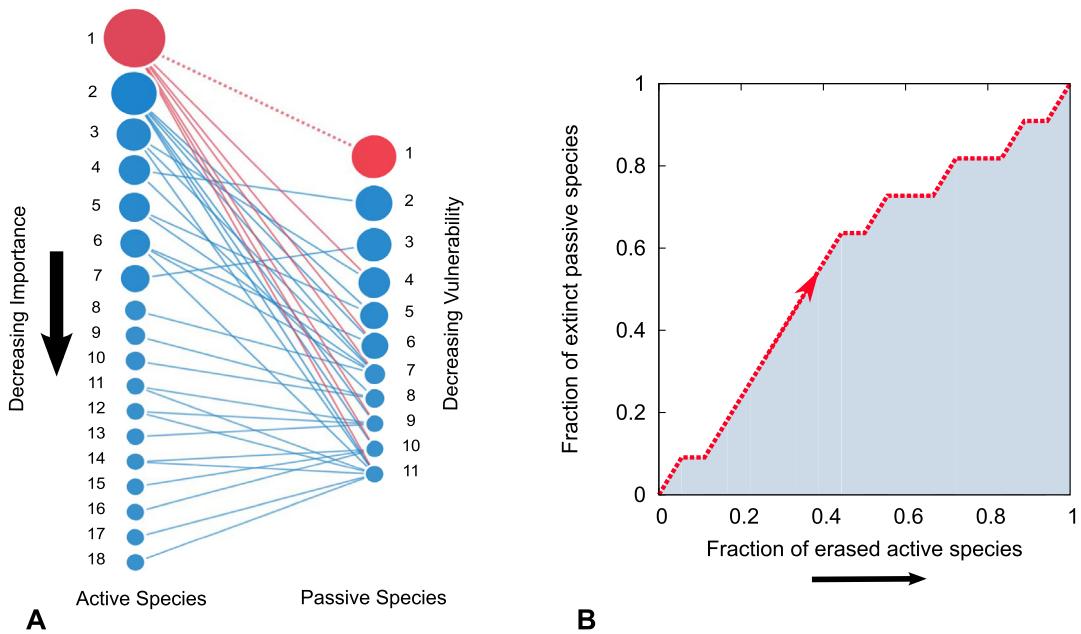


Fig. 25. Illustration of the node removal process (panel A) used to calculate the extinction area metric (panel B). From the bipartite network that connects active and passive species, active species are sequentially removed in order of decreasing “importance” as determined by a given ranking algorithm (panel A). One records the size of co-extinction cascades as a function of the fraction of removed species; the area under this curve is the extinction area of the ranking algorithm. The larger the algorithm’s extinction area, the more accurate the algorithm is in identifying the structurally important active species. One can define a specular extinction area based on a process where passive species are progressively removed in order of increasing vulnerability [71]. The same procedure can be applied to the country–product bipartite network by interpreting countries and products as active and passive nodes, respectively [72].

Source: Reprinted from [71].

design node removal strategies that are based on node centrality metrics other than degree. Crucially, the sequential deletion of the nodes according to different centrality metrics lead to different co-extinction patterns [71,72,335].

An ideal algorithm would provide us with an accurate ranking of the active (passive) species by their importance (vulnerability) in the system [71]. In this respect, we would expect an accurate algorithm to rank first (last) the active (passive) species that cause the biggest damage when removed. The rationale behind this assumption is that an important active species interacts with many passive species, from the generalist to the specialist ones; hence, its removal is likely to cause the secondary extinction of some of the specialist passive species. The main goal of this Section is to present the obtained results on the performance of different ranking algorithms with respect to this benchmark.

Benchmarking ranking algorithms: extinction areas (EAs). The extinction area [335] quantifies the ability of a ranking algorithm to produce a ranking of the nodes that leads to a fast network fragmentation when the nodes are deleted in order of ranking. The EA is closely related to the ATC introduced above. When the active (passive) nodes are progressively removed from the system, one defines the extinction curve (EC) as the fraction of extinct passive (active) nodes as a function of the fraction of removed nodes. The EA is defined as the area under the EC, and it lies within the range (0, 1). The larger the EA, the faster the network breakdown due to the node removal process – see Fig. 25 for an illustration. As emphasized above, the EA critically depends on the removal strategy; we consider two removal strategies:

- Targeting active nodes. We remove the active species in order of decreasing score as determined by a given ranking algorithm.
- Targeting passive nodes. We remove the passive species in order of increasing score as determined by a given ranking algorithm.

We expect a good metric for active species’ importance and passive species’ vulnerability to maximize the respective EA. We stress that while all traditional centralities produce the same ranking in perfectly nested networks [309], the problem becomes non-trivial as soon as imperfect nested structures are considered, such as those observed in empirical datasets.

Results on mutualistic networks. Following Allesina and Pascual [335], two recent works [71,72] have compared different ranking algorithms with respect to their extinction areas in bipartite networks. Inspired by the literature on economic complexity [24], Domínguez-García and Muñoz [71] noticed that different ranking algorithms produce different shapes for the adjacency matrix when the matrix’s rows and columns are ordered by ranking (see our previous discussion in Section 3.3).

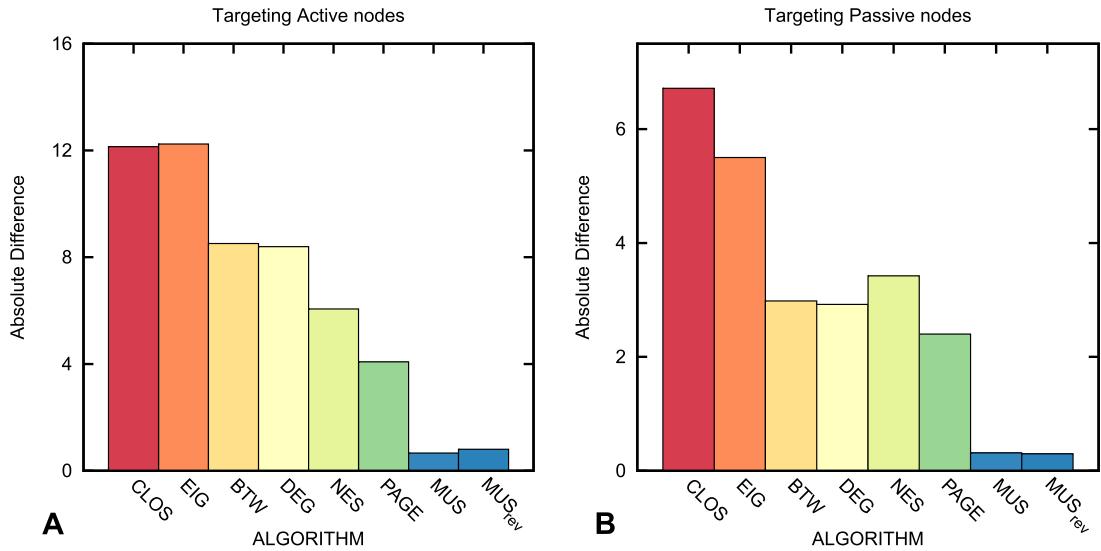


Fig. 26. The ability of different ranking algorithms to quantify node importance for co-extinction cascades as quantified by the extinction area of the co-extinction cascades that they trigger. The Figure shows the difference between the optimal extinction areas obtained by a genetic algorithm and the extinction areas attained by the algorithms, for attacks targeting active (left panel) and passive species (right panel). The metrics included in the analysis are: closeness centrality (*CLOS*), eigenvector centrality (*EIG*), betweenness centrality (*BTW*), degree centrality (*DEG*), inverse contributions to nestedness (*NES*), Google's PageRank (*PAGE*), fitness-complexity algorithm (*MUS*), "reversed" fitness-complexity algorithm (*MUS_{rev}*). We refer to Appendix B for a detailed description of the metrics. The fitness-complexity algorithm gets substantially closer to the maximal extinction area than all the other metrics.

Source: Reprinted from [71].

One can expect ranking algorithms that lead to more nested arrangements of the adjacency matrix to exhibit larger extinction areas than other ranking algorithms. Indeed, as soon as we remove the first few top-ranked nodes from a highly nested structure, the most specialist nodes are likely to become extinct as a result. The same is not guaranteed for a ranking algorithm that does not lead to a packed adjacency matrix.

In line with this expectation, Domínguez-García and Muñoz [71] found that in mutualistic networks, the fitness-complexity metric exhibits systematically larger extinction areas than all the other metrics, including closeness centrality, eigenvector centrality, betweenness centrality, degree centrality inverse contributions to nestedness, Google's PageRank – see Fig. 25 for the main result, and Appendix B for the definitions of the existing metrics. The fitness-complexity metric achieves a similar performance as a computationally expensive genetic algorithm specifically designed by the authors in order to find a ranking of the nodes that maximizes the extinction area. Domínguez-García and Muñoz [71] concluded that the fitness-complexity algorithm should be used to construct rankings of species by importance in mutualistic networks.

Results on country-product export networks. The ranking evaluation procedure described above with the terminology of mutualistic networks can be applied to the country-product export network as well. In this case, the countries (products) play the role of active (passive) species [72]. Mariani et al. [72] found that in world trade networks, the fitness-complexity metric exhibit larger extinction areas than the degree centrality and the method of reflections (a linear ranking algorithm introduced by Hidalgo and Hausmann [110] with the goal to quantify the economic complexity of countries and products). Besides, generalizations of the fitness-complexity metric that lead to more nested matrices exhibit even larger extinction areas. This confirms that non-linear ranking algorithms are the most effective ones in ranking the nodes by their structural importance for co-extinction cascades in bipartite networks.

By exploiting the analogy with ecological networks, in the same way as vulnerable passive species only interact with important active species, "vulnerable" (less complex) products are only exported by "important" (diversified) countries. The fitness-complexity algorithm and its generalizations naturally incorporate this property through the non-linear dependence of product score on country score: indeed, if a product is exported by a low-fitness country, it receives a high penalization by the algorithm (see Eq. (40) and related discussion). This is desirable inasmuch high-complexity products are more likely to be affected by the extinction of high-fitness countries. Non-linear algorithms, therefore, emerge as natural methods to rank the nodes by importance in both mutualistic networks [71] and country-product networks [72].

6.2. Stability and feasibility of systems with competition and mutualism in the population dynamics approach

Understanding how network topology and dynamics affect systemic stability is a key problem in network science [336]. Following the observation that the structure of ecological interaction networks is highly non-random [23], scholars

have sought to understand how the topology of interaction networks impacts the co-existence of species in the system (“feasibility”) and the ability of the system to return to its equilibrium state after a perturbation (“stability”). In absence of empirical temporal data, this impact can be evaluated through numerical simulations or analytic insights based on population dynamics models. To this end, it is essential to introduce a model that incorporates the key elements of mutualistic systems: competition, mutualism, finite interaction handling time [50,337]. Such a model is introduced in Section 6.2.3.

To properly assess the impact of topology on the stable co-existence of species, *structural stability* is a key notion. According to the structural-stability paradigm, when analyzing the outcomes of a population dynamics model, we need to determine not whether nestedness favors biodiversity for a given parameterization of the dynamics, but to determine which topological pattern leads to the stable co-existence of species for the largest region of parameter values [38]. Considering only specific, arbitrary parameterizations of the dynamics can indeed lead to contradictory conclusions on the impact of nestedness on biodiversity [38,325,338,339] – we refer to [38,340] for detailed discussions of this point.

In this Section, we start by reviewing the different notions of local and global stability for a given dynamical system, together with the notion of feasibility (Section 6.2.1). We discuss the global and feasible equilibrium points of interacting systems that feature both competitive and mutualistic interactions, and the impact of nestedness on the range of dynamics parameter that can lead to a stable and feasible equilibrium (Section 6.2.3).

6.2.1. Stability and feasibility of equilibrium points

The first studies on the stability theory of ordinary differential equations date back to the early XX century, with the seminal works by Poincaré and Lyapunov [341]. A fundamental problem in dynamical systems theory is to understand the late-time behavior of a given system. Among all the possible types of asymptotic behavior for a given dynamical system, the simplest one is the convergence toward an equilibrium point. In the following, we review the notion of local and global stability for an equilibrium point.

We narrow our focus to dynamic models described by a set of coupled ODEs in the form

$$\frac{dN_i}{dt} = N_i f_i(\mathbf{N}), \quad (80)$$

where $\mathbf{N} = \{N_1, \dots, N_S\}$ is the set of population sizes, S is the number of species, and $f_i(\mathbf{N}) = N_i^{-1} (dN_i/dt)$ represents species i 's per-capita growth rate. The functions f_i typically depend on a set of parameters that determine the interactions between species. For example, we will sometimes consider a linear per-capita growth rate in the form $f_i(\mathbf{N}) = \alpha_i - \sum_j \beta_{ij} N_j$, where α_i represents the species' *intrinsic growth rates* and $\{\beta_{ij}\}$ represent the interaction coefficients. Such dynamics represents a generalization to N species of the equations originally proposed independently by Lotka [342] and Volterra [343] to study the competitive dynamics of a system composed of two species. Despite its apparent simplicity, this dynamical system can exhibit limit cycle periodic behavior and chaotic behavior conditional on the parameters and the number of included species.³³

In the rest of this Section, we introduce three properties of equilibrium points: local stability, global stability, and feasibility. Importantly, insights from random matrix theory on the persistence of ecologic communities (see Section 6.3) are mostly based on the local stability [37,326] and, recently, feasibility [345,346]. Recent insights from population dynamics [38], see Section 6.2 focused on globally stable and feasible equilibrium points.

Local stability. A given equilibrium point \mathbf{N}^* is *locally stable* if, after a small external perturbation, the system returns to the equilibrium point [341]. To determine whether a given equilibrium point \mathbf{N}^* is locally stable or not, one needs to find the eigenvalues of the so-called “community matrix” [347] whose elements are defined as

$$M_{ij}(\mathbf{N}^*) = \left. \frac{\partial (N_i f_i(\mathbf{N}))}{\partial N_j} \right|_{\mathbf{N}=\mathbf{N}^*}. \quad (81)$$

The community matrix corresponds to the system's Jacobian matrix evaluated at the equilibrium point \mathbf{N}^* . A given equilibrium point is locally stable if and only if the real parts of the eigenvalues of $M_{ij}(\mathbf{N}^*)$ are all negative. For a linear per-capita growth rate in the form $f_i(\mathbf{N}) = \alpha_i - \sum_j \beta_{ij} N_j$, we obtain $M_{ij} = -N_i^* \beta_{ij}$ for each equilibrium point \mathbf{N}^* [348]; such equilibrium point is locally stable if $\beta_{ij} \geq 0$ for all i, j .

Feasibility and global stability. A given equilibrium point \mathbf{N}^* is *feasible* if $f(\mathbf{N}^*) = 0$ and $N_i^* > 0$ for all species i [38,348,349]. A feasible equilibrium point describes, therefore, a scenario where all species can co-exist. A given equilibrium point \mathbf{N}^* is *globally stable* if \mathbf{N}^* is a global attractor, and the trajectories of the dynamic system converge to \mathbf{N}^* regardless of the initial condition [38,348]. The fundamental difference between local and global stability is that while local stability studies whether a system will return to the original equilibrium point after an infinitesimally small perturbation, global stability studies whether the system will return to the original equilibrium point after perturbations of any magnitude [38]. Is a given equilibrium point globally stable? In general, no: we will provide a simple example of an equilibrium point that is not globally stable in Section 6.2.2. Linear algebra theorems nevertheless provide us with sufficient conditions for an equilibrium point to be stable. It is convenient to introduce the following definitions:

³³ We refer the interested reader to the books [341,344] for details.

- A matrix M is *positive-defined* if the eigenvalues of $M + M^T$ are all positive.
- A matrix M is *Volterra-dissipative* if and only if there exists a diagonal matrix D such that $DM + M^T D$ is positive-defined.

The following theorems hold [348]:

- If a matrix is positive-defined, then it is Volterra-dissipative.
- For a model with linear per-capita growth rate, $f_i(\mathbf{N}) = \alpha_i - \sum_j \beta_{ij} N_j$, we can arrange the interaction coefficients in a $S \times S$ matrix, which we denote as B . If B is Volterra-dissipative, then (1) Any feasible equilibrium is globally stable; (2) There exists one unique globally stable equilibrium point.³⁴

The previous theorems imply that if the matrix of interaction strengths is positive-defined, then the system has a unique globally stable equilibrium. In Section 6.2.3 we shall see that, in models with competition and mutualism, this property can be used to set an upper bound to the level of mutualistic strength up to which global stability is guaranteed.

6.2.2. Stability and feasibility in a two-species competition model

To get an intuition about the implications of the different stability definitions and of the feasibility condition, we consider here a simple two-species competition model. The proposed example and the calculations below closely follow those presented in [38]. The system evolves according to the generalized Lotka–Volterra equations:

$$\begin{cases} \frac{dN_1}{dt} = N_1 (\alpha_1 - \beta_{11} N_1 - \beta_{12} N_2) \\ \frac{dN_2}{dt} = N_2 (\alpha_2 - \beta_{21} N_1 - \beta_{22} N_2) \end{cases} \quad (82)$$

In line with the notation of Eq. (80), the previous equations can be written in the compact form

$$\frac{d\mathbf{N}}{dt} = \mathbf{N} \mathbf{f}_{\alpha, B}(\mathbf{N}), \quad (83)$$

where we defined $\mathbf{f}(\mathbf{N}) = (f_1(\mathbf{N}), f_2(\mathbf{N}))$, $f_1(N_1, N_2) = \alpha_1 - \beta_{11} N_1 - \beta_{12} N_2$, $f_2(N_1, N_2) = \alpha_2 - \beta_{21} N_1 - \beta_{22} N_2$. It is also convenient to introduce the matrix B whose elements are the interaction coefficients β_{ij} . The equilibrium point (N_1^*, N_2^*) is found by imposing $f_1(N_1^*, N_2^*) = 0$ and $f_2(N_1^*, N_2^*) = 0$, which has the solution

$$\begin{cases} N_1^* = \frac{\alpha_1 \beta_{22} - \alpha_2 \beta_{12}}{\beta_{11} \beta_{22} - \beta_{21} \beta_{12}} \\ N_2^* = \frac{\alpha_2 \beta_{11} - \alpha_1 \beta_{21}}{\beta_{11} \beta_{22} - \beta_{21} \beta_{12}} \end{cases} \quad (84)$$

According to the theorem presented in Section 6.2.1, if the matrix B is positive-defined, then this equilibrium point is globally stable. Therefore, we narrow our attention to $\{\beta_{ij}\}$ values such that $\beta_{11} \beta_{22} - \beta_{21} \beta_{12} > 0$. Graphically, the equilibrium point (N_1^*, N_2^*) corresponds to the intersection point between the isolines $f_1(N_1, N_2) = 0$ and $f_2(N_1, N_2) = 0$ in the (N_1, N_2) plane (see Fig. 27A–C). The equilibrium point is feasible if the intersection point lies in the quadrant with $N_1 > 0$ and $N_2 > 0$. Importantly, different settings of the model parameters r_1, r_2 , lead to different isolines in the (N_1, N_2) plane; some parameter settings may lead to a feasible equilibrium point (see Fig. 27A), whereas other parameter settings may not (see Fig. 27B–C).

The equilibrium point (N_1^*, N_2^*) is feasible if and only if both $N_1^* > 0$ and $N_2^* > 0$, which leads to the conditions :

$$\begin{cases} \alpha_1 \beta_{22} - \alpha_2 \beta_{12} > 0 \\ \alpha_2 \beta_{11} - \alpha_1 \beta_{21} > 0 \end{cases} \quad (85)$$

If all the parameters $\{\beta_{ij}\}$ are positive, we can compress the two previous conditions into a single condition for the ratio r_1/r_2 between the intrinsic growth rates:

$$\frac{\beta_{12}}{\beta_{22}} < \frac{\alpha_1}{\alpha_2} < \frac{\beta_{11}}{\beta_{21}}. \quad (86)$$

Such condition is verified in the region of the (α_1, α_2) plane bounded by the two straight lines $\alpha_2 = (\beta_{12}/\beta_{22})\alpha_1$ and $\alpha_2 = (\beta_{11}/\beta_{21})\alpha_1$.

The results on the simple two-species competition model described by Eq. (82) shed light on two fundamental facts [38]:

- The feasibility of a given equilibrium point critically depends on the dynamics parameters. Some sets of dynamics parameters may lead to a feasible equilibrium point (such as in Fig. 27A) or it may not (such as in Fig. 27B–C).
- One can identify a "feasibility domain" in the intrinsic growth rates' plane where the dynamics' equilibrium point is feasible (see Fig. 27D). Crucially, the size of such domain depends on the dynamics parameters.

³⁴ Such equilibrium point is not necessarily feasible, as some of the population sizes may be equal to zero. See the next paragraph for the definition of feasible equilibrium point.

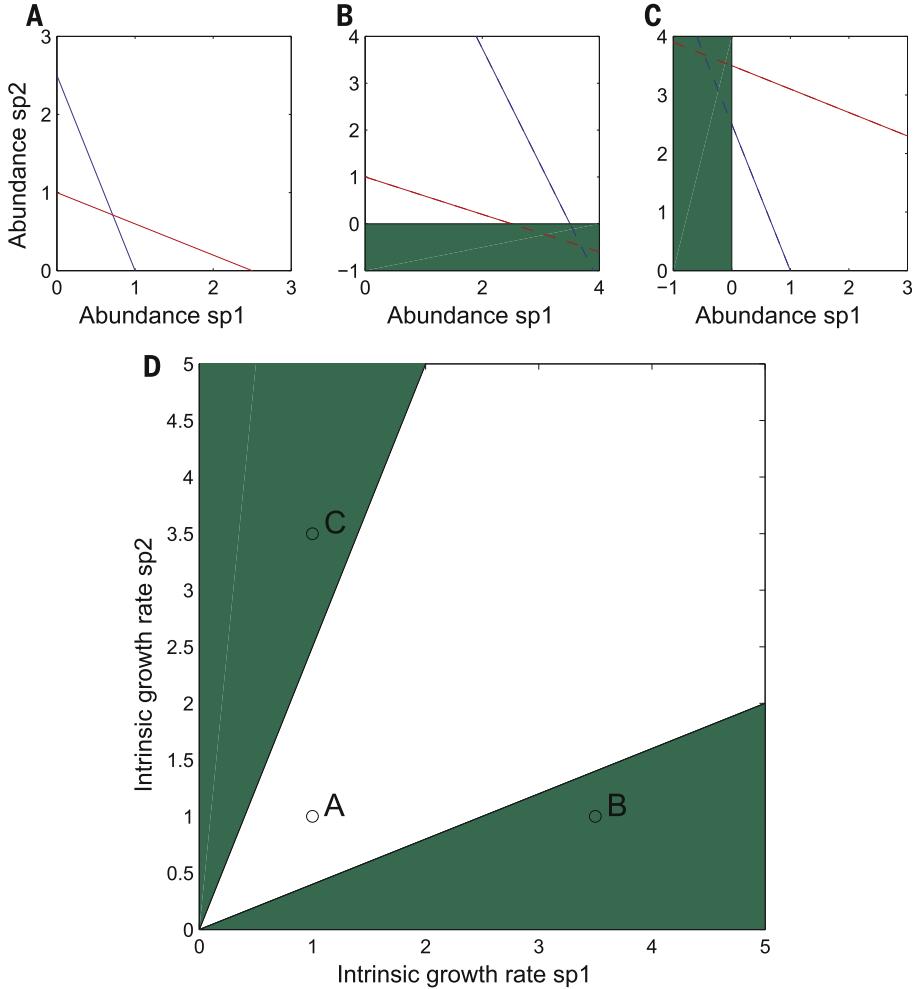


Fig. 27. Global stability and feasibility of the equilibrium points of the simple mutualistic dynamics described by Eq. (82). Panels A to C represent the isoclines $f_1(N_1, N_2) = 0$ and $f_2(N_1, N_2) = 0$ for three different parametrizations of the intrinsic growth rates (α_1, α_2) ; their intersection point is the equilibrium point of the dynamics. Among the three scenarios A–C, only A represents a feasible equilibrium point ($N_1 > 0$ and $N_2 > 0$). Panel D shows that the intrinsic growth rates space can be partitioned into regions where the equilibrium is feasible (white area), and regions where the equilibrium is not feasible (green areas). For this simple dynamics, the feasibility domain is given by Eq. (86). The three points A, B, C correspond to the three above panels; their position in the intrinsic growth rates space confirms that only A is feasible.

Source: Reprinted from [38].

Section 6.2.3 shows that these properties also hold for a mutualistic model of interaction. Based on these properties, Rohr et al. [38] put forward the notion of *structural stability*: when studying the dynamical stability of a given ecological system, the question is not whether the dynamics' equilibrium point is feasible, given a set of dynamics parameters. The main question is whether the parameter region composed of the parameters that lead to a feasible equilibrium point is large, and how the size of such region depends on network topology [38]. Next Sections will investigate this fundamental question for a model with competition and mutualism, and reveal the strong dependence of its answer on the level of nestedness in the network.

In a similar spirit to the simple two-species described above, Bastolla et al. [350,351] analyzed more complex competitive systems, and derived analytical conditions that must be fulfilled by the moments of the growth rate distribution in order for the system to be feasible. The framework introduced by Bastolla et al. [350,351] has been later extended [50] to systems with both competition and mutualism. We refer to [340] for a detailed presentation of the results in [350,351].

6.2.3. Stability and feasibility in models with competition and mutualism

The previous Section considered a simple, two-species competition model. Following Rohr et al. [38], we extend here those insights to the N -species model with (inter- and intra-specific) competition and mutualism studied by Bastolla

et al. [50]. The model incorporates network topology, interspecific competition, and mutualistic interactions with finite handling time (see below). To introduce the main elements of the model, we start from a simpler two-species model [352], and later describe the full N -species model [50].

A two-species model with intraspecific competition and mutualism. Before describing the results for an ecosystem composed of multiple species, let us consider a simple two-species world where the two species represent animals and plants, respectively. Our simplified world features both intraspecific competitive interactions and mutualistic interactions between plants and animals. For an isolated “animal” species whose A individuals compete for the same resource, it is natural to consider a logistic population model

$$\frac{dA}{dt} = A(\alpha - \beta A), \quad (87)$$

where α is the species’ intrinsic per-capita growth rate, and β is a parameter that quantifies the strength of intra-species competitive interactions. To simplify the discussion, we refer to the individuals of this species as “animals”.

Mutualistic interactions can be incorporated by introducing a second species, whose number of individuals is denoted as P – we refer to the individuals of this species as “plants”. The interactions between animals and plants are mutualistic, which results in a positive contribution to dA/dt from the plants. It is plausible to assume the benefits for animals from mutualistic interactions to increase with the number P of plants. However, Wright [352] noticed that this contribution cannot be unbounded. Let us consider the situation where the animals are bees and the plants are the flowers that they pollinate. A bee needs a certain amount of time to travel to the plant, collect enough pollen from a flower, and travel back to its nest to provision its larval cells. Such time – referred to as *handling time* [352] – limits the benefits that the animal species can obtain from its mutualistic interaction with the plant species.

To incorporate both handling time and mutualistic interactions, Wright [352] assumed a dynamic equation of the form

$$\frac{dA}{dt} = A \left(\alpha - \beta A + \frac{c \gamma P}{1 + h \gamma P} \right). \quad (88)$$

The form of the mutualistic interaction term $c \gamma P/(1 + h \gamma P)$ is in line with the empirical observations by Holling on the capability of a consumer to process a given amount of food resource [353]. The non-linear functional response in the form $\gamma P/(1 + h \gamma P)$ is also referred to as Type II response [353]. Importantly, if $h = 0$, the contribution from mutualistic interactions is unbounded. By contrast, for any $h > 0$, the mutualistic term is bounded by h^{-1} , i.e., by the value it assumes in a world with a large number $P \gg 1$ of plants or, equivalently, in a world with strong mutualistic interactions ($\gamma \gg 1$). In a purely-mutualistic world with $\alpha = 0$, $\beta = 0$, and $\gamma \gg 1$, the parameter h determines the rate at which mutualistic interactions are converted into new animals: $dA = A c (dt/h)$. The model described by Eq. (88) constitutes the basis for the model with competition and mutualism studied by Bastolla et al. [50].

Generalizing the model to an arbitrary number of species. Different types of N -species mutualistic population-dynamics models have been studied in the literature. A model without handling time for a fully mixed population (without network structure) was considered by Bascompte et al. [354] to analytically derive conditions for a system to be feasible and stable. Commenting on [354], Holland et al. [355] pointed out the necessity of including the saturation of the mutual benefit two species can gain from their interaction, an aspect that was already stressed by May [356] and included in Wright’s model [352] presented above. Okuyama et al. [337] considered a model with mutualistic interactions, intraspecific competitive interactions, network structure, and finite handling time.

Here, we present the version of the model used by [38,50]. In the model, there are two classes of species: plants and animals. We assume that there are N animal species and M plant species whose populations are denoted as $\mathbf{A} = \{A_1, \dots, A_N\}$ and $\mathbf{P} = \{P_1, \dots, P_M\}$, respectively. The dynamical equations for the evolution of the species’ populations are [38,50]

$$\begin{cases} \frac{dP_i}{dt} = P_i \left(\alpha_i^{(P)} - \sum_j \beta_{ij}^{(P)} P_j + \frac{\sum_j \gamma_{ij}^{(P)} A_j}{1 + h \sum_j \gamma_{ij}^{(P)} A_j} \right) \\ \frac{dA_i}{dt} = A_i \left(\alpha_i^{(A)} - \sum_j \beta_{ij}^{(A)} A_j + \frac{\sum_j \gamma_{ij}^{(A)} P_j}{1 + h \sum_j \gamma_{ij}^{(A)} P_j} \right) \end{cases} \quad (89)$$

where $\{\alpha_i^{(A)}, \alpha_i^{(P)}\}$ represent the intrinsic growth rates, $\{\beta_{ij}^{(A)}, \beta_{ij}^{(P)}\}$ are the parameters that quantify the strength of competitive interactions between species of the same kind; $\{\gamma_{ij}^{(A)}, \gamma_{ij}^{(P)}\}$ are the parameters that quantify the strength of beneficial (mutualistic) interactions between animal and plants; h is the *handling time* [352] which implies an upper bound to the benefits a species can receive through interaction with species of different type.³⁵ Recent works [38,357] further parameterized the mutualistic interaction coefficients γ_{ij} as

$$\gamma_{ij}(\delta) = \gamma_0 \frac{A_j}{k_i^\delta}, \quad (90)$$

³⁵ The handling time here plays the same role as in Eq. (88).

where γ_0 represents the mutualistic strength, and δ is a parameter called *mutualistic trade-off* [357]. The mutualistic trade-off factor $k_i^{-\delta}$ reduces the benefit from mutualistic interactions with species that have many interactors, which is justified by field observations [358].

Crucially, the model defined by Eq. (89) features interspecific competitive interactions. To keep the competitive interactions as simple as possible, some works [38,50] adopted a mean-field description of the interspecific competition by posing $\beta_{ii}^{(P)} = \beta_{ii}^{(A)} = 1$ (*intraspecific competition*), and $\beta_{ij}^{(P)} = \beta_{ij}^{(A)} = \lambda < 1$ for $i \neq j$ (*interspecific competition*). Essentially, species of a type are assumed to compete with equal strength with all the species of the same type. The mean-field assumption can be relaxed: for example, Gracia-Lázaro et al. [359] studied a similar metapopulation model, but they assumed that species do compete more intensely with species that are interested in the same resources. More specifically, they assumed that a pairwise competition term between two plant species i and j proportional to the number of common interaction partners of i and j .

Variants of this population dynamics model and alternative models can be also relevant. Besides competitive and bipartite mutualistic interactions, we notice that recent field experiments [360] indicate that plant–plant facilitation interactions can have a substantial impact on the structure and robustness of pollination networks, which may motivate the inclusion of these interactions in population-dynamics models. One can also consider population dynamics models where the foraging effort of a given pollinator on a given plant is a heterogeneous and adaptive variable [361]. The presence of adaptive foraging can substantially alter the impact of nestedness on species persistence – we refer the interested reader to [362] for the details.

Preliminary results on the model with competition and mutualism. The model defined by Eq. (89) and its variants have been used in several works to assess how specific network structures affect the co-existence of species in mutualistic and trophic networks. More specifically, early works based on the population dynamics approach focused on the question: given a parameterization of the intrinsic growth rates and the interaction parameters, how does a given network structure impact on the properties of the dynamics' equilibrium point? We mention here some of the most prominent results obtained with the goal of addressing this research question.

Okuyama and Holland [337] considered the model without interspecific competition (i.e., in the notation of the previous paragraph, $\lambda = 0$), and they assumed that the degree distribution p_k of mutualistic communities follows a power-law $p_k = C k^{-\gamma}$. They found that the resilience of mutualistic communities only depends on nestedness when the power-law exponent γ is large enough; in this case, resilience correlates positively with nestedness. More recently, Morone et al. [363] found that the model without interspecific competition features a sharp transition from a “feasible” phase (where the average number of individuals per species is larger than zero) to a “collapsed” phase (where the average number of individuals per species is equal to zero). The critical point that separates the two phases such transition can be analytically estimated based on the maximal k -core of the network [363].

Bastolla et al. [50] used the model with interspecific and competitive interaction to show both analytically and numerically that in real mutualistic networks, nestedness has a positive impact on the system's biodiversity increase due to mutualism. More specifically, the analytic framework developed by Bastolla et al. [50] clarified that in a competitive system, introducing mutualism alters the “effective competition” between species; crucially, the potential mitigation of competitive effects is modulated by network topology. The resulting increase of biodiversity due to mutualism can be quantified as the derivative of the predicted maximum number of species with respect to the mutualism-to-competition ratio R , calculated at $R = 0$ – see Eq. (6) in [50]. The mutualism-to-competition ratio R is expressed in terms of the model's parameters, and it grows as the relative strength of mutualistic interaction increases; $R = 0$ corresponds to the absence of mutualism. Importantly, nestedness is positively correlated with the so-defined increase of biodiversity, which led Bastolla et al. [50] to conclude that “nestedness reduces effective interspecific competition and enhances the number of coexisting species”.

As the analytic theory developed in [50] has been already detailedly reviewed in Appendix G in [9] and in [340], we will not present its details here, and focus on the framework developed by Rohr et al. [38] in the following. It is worth mentioning that the conclusion that nestedness increases the biodiversity was later challenged by James et al. [325] in a paper that has been subsequently debated [338,339]; however, the conclusions by [325] are based on a specific choice of the parameters of the dynamics, and they are based on an incorrect formula for the predicted maximum number of species in the community [340].

Thébault and Fontaine [324] consider a variant of the dynamics described by Eq. (89) that can describe both mutualistic and trophic interaction. The core difference between mutualistic and trophic interactions is the sign of the term that represents animal–plant interactions. They found that network structure has an opposite effect on the stability of mutualistic networks as compared to the stability of trophic networks. In particular, nestedness has a positive effect on the resilience of mutualistic networks, defined as the speed at which the community returns to equilibrium after a perturbation in the system. By contrast, it has a strongly negative effect on the persistence of trophic networks, defined as the proportion of existing species after equilibrium is reached.

Rohr et al. [38] emphasized that when we fix the model parameters and we find a feasible fixed point, (1) it is not guaranteed that the found equilibrium point is globally stable, i.e., that the system's dynamics will converge toward that state for any given initial condition; (2) it is not guaranteed that we will find the found equilibrium point for other parameterizations of the intrinsic growth rates. Besides, for a given system and two given network topologies (e.g., a random and a nested one), it is possible to find a set of intrinsic growth rates such that one of the two is persistent and the other one is not, and another set such that the opposite holds – see Fig. 2 in [38] for an illustration.

Domain of feasibility in the mutualistic model. The notion of *structural stability* has been long-known in developmental biology [364] and engineering [365] – we refer to [340] for a historical overview. Early precursors applied the notion of structural stability to ecological models of a few interacting species [366–368]. More recently, structural stability has been applied by Bastolla et al. [350,351] for purely competitive systems, and by Bastolla et al. [50] for systems with competition and mutualism.

The key idea behind structural stability analysis in ecology is that when studying a population dynamics, we should not pre-impose an arbitrary parameterization of the intrinsic growth rates, but study a large portion of possible growth rate values, and quantify the size of the region of the growth rate space – called *feasibility domain* – where the system reaches a feasible equilibrium point [38]. For the two-species competition model described in Section 6.2.2, the feasibility domain can be determined analytically (see Eq. (86) and Fig. 27). A similar derivation for the mutualistic model defined by Eq. (89) is not possible. An exhaustive numerical search is also impossible due to the large size of the parameter space of intrinsic growth rates.

To determine the size of the feasibility domain, there are at least two approaches, a geometrical and a probabilistic one. In the geometric approach, for a linear dynamics of the form $\dot{\mathbf{N}} = -\mathbf{N}(\boldsymbol{\alpha} + \mathbf{B}\mathbf{N})$ ruled by an interaction matrix \mathbf{B} , the feasibility domain is defined as [369]

$$D_F(\mathbf{B}) = \{\boldsymbol{\alpha} = -\mathbf{B}^{-1}\mathbf{N}^* \mid N_i^* > 0 \quad \forall i\} \quad (91)$$

By using expressions for solid angles in dimension larger than three [370], one can compute the volume $\Omega(\mathbf{B})$ of $D_F(\mathbf{B})$ as [320,369]

$$\Omega(\mathbf{B}) = \frac{1}{(2\pi)^S/2 \sqrt{|\det(\mathbf{B})|}} \int \dots \int_{\mathbf{N}^* > 0} \exp\left(-\frac{1}{2}\mathbf{N}^{*T} \mathbf{B}^T \mathbf{B} \mathbf{N}^*\right) d\mathbf{N}^*; \quad (92)$$

we refer to [369] for the details, and to [371] for the generalization to non-linear population dynamics models. A simpler approach to estimate the size of the feasibility domain is to compute the average probability $\omega(\mathbf{B})$ that a randomly selected species i is feasible, i.e., $N_i^* > 0$ [38,369]. In order for an equilibrium point to be feasible, all species must be feasible; therefore, for sufficiently weak interactions, the feasibility domain size $\Omega(\mathbf{B})$ can be simply estimated as $\Omega(\mathbf{B}) \simeq (\omega(\mathbf{B}))^S$ [369]. This approach has been adopted by Rohr et al. [38] to assess the impact of nestedness on the feasibility domain size (see Section 6.2.4).

Critical value of mutualism for global stability in the mutualistic model. A feasible equilibrium point is not necessarily globally stable, and a globally stable equilibrium point is not necessarily feasible. The existence of a unique globally stable equilibrium point is only guaranteed when the interaction matrix \mathbf{B} is positive-defined, i.e., when all its eigenvalues have a positive real part (see Section 6.2.1). Numerically, one finds that in the linearized version of the model with competition and mutualism (i.e., the model defined by Eq. (89) with handling times equal to zero), the real part of the largest eigenvalue of \mathbf{B} tends to decrease with the mean mutualistic interaction strength $\bar{\gamma}$. Hence, the convergence to a unique globally stable equilibrium is guaranteed up to a critical value $\hat{\gamma}$ of the interaction strength; such point is numerically determined as the smallest $\bar{\gamma}$ value such that one of the eigenvalues of \mathbf{B} reaches zero [38,320]. The larger $\hat{\gamma}$, the stronger mutualistic interactions can be tolerated by the system without losing global stability, i.e., the more globally stable the system [320]. Therefore, one can study the impact of a structural property on the global stability of the system by studying its relation with $\hat{\gamma}$: a positive correlation would imply that structural property favors the global stability of the system.³⁶

6.2.4. The impact of nestedness on feasibility and global stability

The structural stability framework allows us to address one of our original questions: what is the impact of nestedness on the feasibility of a given system? Based on our above considerations, to address this question for a given empirical network, one can generate model networks with different level of nestedness through a resampling procedure that preserves the total number of species and the expected number of interactions, and then study how the probability ω that a randomly-selected species is feasible is affected by nestedness [38]. This probability ω tends to be smaller the closer we are to the boundaries of the feasibility domain. To control for this possible confounding factor, first, Rohr et al. [38] determined the center $\{\boldsymbol{\alpha}_S^{(A)}, \boldsymbol{\alpha}_S^{(P)}\}$ of the feasibility domain. This vector is the vector that can tolerate the maximal perturbation before leaving the feasibility domain, and it can be determined analytically – see the Supplementary Material of [38] for details. Second, they quantified the deviation of a given set of intrinsic growth rates from the structural vector of intrinsic growth rates. Intuitively, this deviation η (which is an angle in the intrinsic growth-rates space) quantifies how far a set of intrinsic growth rates is from the “center” of the feasibility domain, and it is used as a control variable when assessing the impact of network topology on the feasibility domain size.

Rohr et al. [38] studied the impact of various factors on the fraction of surviving species: deviation from the structural vector of plants η_P and animals η_A , network nestedness \mathcal{N} (as determined by NODF), mutualistic trade-off δ (see Eq. (90))

³⁶ In this discussion, we have referred to the linearized version of the model ($h = 0$). The critical value $\hat{\gamma}$ for the non-linear dynamics is larger than or equal to the critical value for the linearized dynamics, which implies that the critical value $\hat{\gamma}$ determined with the linearized dynamics serves as an upper bound to the range of mutualism that allows for global stability. We refer to [38] for all the details.

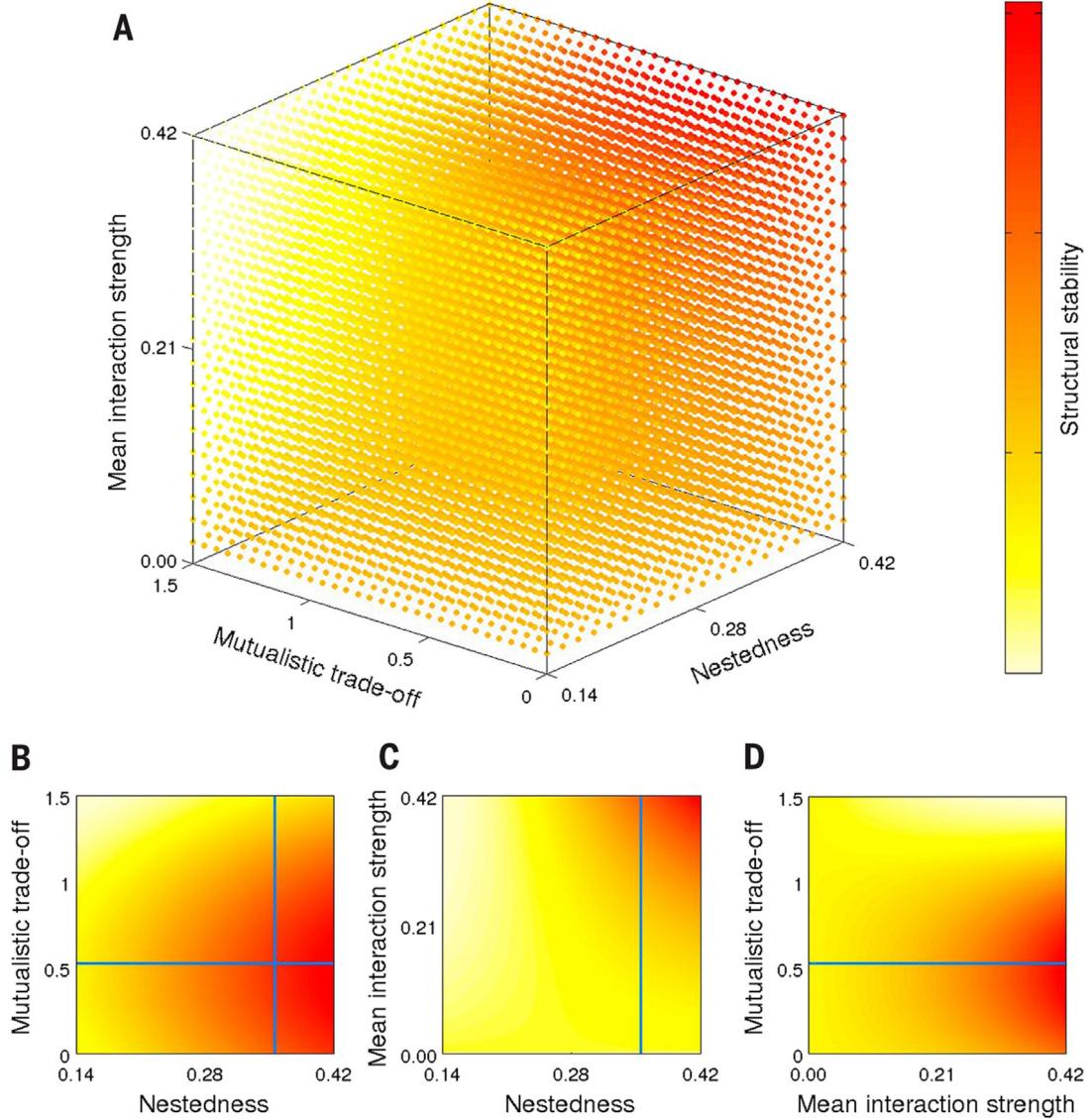


Fig. 28. The overall contribution of nestedness, mutualistic interaction strength, and mutualistic trade-off to the fraction of surviving species. The larger such contribution (determined by Eq. (94)), the larger the region of the intrinsic growth rate space which is compatible with the stable co-existence of species. The results on this empirical network (mutualistic interactions located in the grassland asclepiads in South Africa – results for different datasets are qualitatively similar) show that more stable structures tend to be associated with larger nestedness, larger interaction strength, and smaller mutualistic trade-off. The panels below show two-dimensional slices of the three-dimensional feature space above.
Source: Reprinted from [38].

for the definition) and mean interaction strength $\bar{\gamma}$ ($\bar{\gamma} = \sum_{ij} \gamma_{ij}$). Based on these variables, one can consider a generalized linear model for the probability ω that a randomly selected species i is feasible (i.e., $N_i^* > 0$) defined by the equation [38]

$$\log\left(\frac{\omega}{1-\omega}\right) \sim \log(\eta_A) + \log(\eta_P) + \bar{\gamma} + \bar{\gamma}^2 + \bar{\gamma}N + \bar{\gamma}N^2 + \bar{\gamma}\delta + \bar{\gamma}\delta^2, \quad (93)$$

where the quadratic terms account for possible non-linear effects. The common factor $\bar{\gamma}$ in the contributions of nestedness and mutualistic trade-off reflects the fact that if the mutualistic interaction strength is zero, then nestedness and mutualistic trade-off cannot contribute to the fraction of surviving species. The joint impact of nestedness, mutualistic-trade off, and mutualistic interaction strength on the feasibility domain size is quantified through the partial fitted values

$$SS = \hat{\beta}_1 \bar{b} + \hat{\beta}_2 \bar{b}^2 + \hat{\beta}_3 \bar{b}N + \hat{\beta}_4 \bar{b}N^2 + \hat{\beta}_5 \bar{b}\delta + \hat{\beta}_6 \bar{\gamma}\delta^2. \quad (94)$$

where $\hat{b}_1, \dots, \hat{b}_6$ are the coefficient of the respective terms that maximize the likelihood of the model described by Eq. (93). The larger this contribution is, the larger the size of the feasibility domain as a consequence of these factors.

The results by Rohr et al. [38] (see Fig. 28) show that a larger degree of nestedness tends to be associated with a larger feasibility domain. The most structurally stable systems are characterized by maximal nestedness, small mutualistic trade-off, high mutualistic-interaction strength (compatible with the constraint that all feasible solutions are globally stable) [38]. Besides, for each empirical network, one can determine the largest theoretical feasibility domain, and compare it with the feasibility domain of the observed networks. The results by Rohr et al. (Fig. 6 in [38]) suggest that the structure of observed empirical networks tends to maximize the size of the feasibility domain.

While these results point out the positive impact of nestedness on the size of the feasibility domain, a larger feasibility domain is not necessarily associated with a more globally stable system. It turns out that while nestedness is indeed associated with larger feasibility domains [38,320], it also tends to decrease the global stability of the system (i.e., the upper value of mutualistic strength γ up to which global stability is guaranteed is negatively correlated with nestedness, see Fig. 2 in [320]). These findings suggest that “nestedness tends to promote feasibility over stability” [320]. Hence, it is natural to ask: does the dynamics of real communities tend to promote feasibility or stability in the long term? To address this question, Saavedra et al. [320] analyzed a time-stamped empirical dataset of an Arctic pollinator community located in Greenland (1996–1997, [372]). Their results indicate that newcomer species in the system tend to promote feasibility and nestedness over stability. Saavedra et al. [320] stress that the assumption that dynamical stability is essential to the persistence of a community might not be justified, and feasibility might be a more fundamental property for species co-existence. We expect additional empirical studies to further address this fundamental point for our understanding of the persistence of ecological communities.

6.3. Local stability of mutualistic systems in the random-matrix theory approach

The results presented in Section 6.2 concerned the impact of nestedness on the size of the parameter regions that are compatible with feasible and globally stable equilibrium points. They were obtained by studying a population dynamics model. The need for a detailed knowledge of the dynamics is bypassed by the random-matrix theory approach. In this alternative approach, one generates a random community matrix M with some given properties, and studies the local stability [37] and feasibility [345] of its equilibrium points. This means that instead of postulating a dynamics and deriving its community matrix M through Eq. (81), one directly generates M and studies its properties.

The advantage of this approach is that it allows us to estimate analytically the stability conditions for a variety of interaction topologies, including a random, a mutualistic, and a nested topology — the corresponding results are presented in Sections 6.3.1–6.3.3, respectively. As this review focuses on nestedness, we will not cover the results for other interaction types (such as predator-prey interactions, competitive matrices, etc.); we refer the interested reader to [37] for these results.

While random-matrix theory allows for various analytic developments, it is important to be aware of its assumptions. Bascompte and Ferrera [340] pointed out that the generality of the results obtained in [37] is limited by two factors: (1) The mutualistic bipartite matrices studied in [37] (see Section 6.3.2) do not feature interspecific competition which is instead a critical component of mutualistic systems [340]; (2) The local-stability criterion adopted in [37] is purely based on the community matrix's eigenvalue λ^M with the largest real part, which provides us with a necessary condition for systemic instability. On the other hand, it is more complex to determine the system's *distance-to-instability* — i.e., the maximal perturbation that the system can absorb before losing local stability — which also depends, in general, on the type of perturbation [340]. Nevertheless, random-matrix theory plays a central role for many analytic results on the impact of complexity and network structure on stability and feasibility [37,326,345], which is why it is important for the reader to understand its rationale and basic elements.

6.3.1. Stability-complexity relation: May's result for random matrices

In the random-matrix theory approach, the first step is the generation of a random $S \times S$ community matrix M , where S is the number of species in the system. Each element M_{ij} of the community matrix is, therefore, a random variable X with mean $\langle X \rangle$ and standard deviation σ . We further denote by C the matrix connectance, and by $-d$ the value of the diagonal elements. The simplest topology is arguably the random one [94]. To generate a random-topology matrix, we extract, for each off-diagonal matrix element M_{ij} ($i \neq j$), a random number p from the uniform distribution $\mathcal{U}[0, 1]$. If $p \leq C$, we extract a random number X from $\mathcal{N}(0, \sigma^2)$, and set $M_{ij} = X$. By contrast, if $p > C$, we set $M_{ij} = 0$. The diagonal terms are fixed: $M_{ii} = -d$ for all species i .

As we have seen in Section 6.2.1, the equilibrium points of the dynamics ruled by a community matrix M is locally stable if the real part of all M 's eigenvalues is negative. One can prove.³⁷ that for the random matrices considered in this Section, this holds with high probability when [37]

$$\sqrt{SC} < \theta = \frac{d}{\sigma}. \quad (95)$$

³⁷ This results is a consequence of Girko's circular law [373,374]. Let us start by setting all the diagonal elements d equal to zero. For any distribution of X with zero mean and finite variance $\sigma^2 < \infty$, in the limit $S \rightarrow \infty$, the eigenvalues of M are then uniformly distributed inside the circle of radius $\sigma\sqrt{S}C$ centered at $(0, 0)$ in the complex plane. Non-zero diagonal elements $-d$ shift the circle's center from $(0, 0)$ to $(-d, 0)$. In order to have all the eigenvalues' real parts smaller than zero, we need the radius of the circle to be smaller than d , i.e., $\sigma\sqrt{S}C < d$.

This result poses a constraint to the number of species to a system that aims to be stable against small external perturbations. If, for instance, $C = 0.1$, $X \sim \mathcal{N}(0, \sigma^2)$, $\theta = 2$, Eq. (95) poses $S < 41$ [37] as the necessary condition for local stability. Besides, Eq. (95) suggests a stability–complexity dualism: complexity (to be interpreted as a large number S of interacting species with a large fraction C of interacting species) begets instability.

Derived in 1972 by Robert May [94], this celebrated result has deeply influenced subsequent developments in theoretical ecology. Subsequent works have explored the stability–complexity relationship in more complex models of interaction [37,375,376] and, more recently, the relation between feasibility and complexity [345]. For example, Allesina and Tang have generalized these relations to competitive, mutualistic, predator–prey, and mixed interactions; Stone [376] has generalized them to competition matrices where each element has an additional competition term that is uniform across the nodes, and to mutualistic systems. Reviewing the literature on the stability–complexity relationship falls out of the scope of this article; in the following, we focus on the impact of interaction topology on the stability–complexity relationship for bipartite mutualistic interactions.

6.3.2. Stability of mutualistic networks

The random matrices by May are highly unrealistic: each pair of species interacts with the same probability. For sufficiently large matrices, this randomness results in fixed frequencies for each type of interaction: for example, predator–prey interactions ($M_{ij} > 0$ and $M_{ji} < 0$) are twice as frequent as mutualistic ($M_{ij} > 0$ and $M_{ji} > 0$) and competitive ($M_{ij} < 0$ and $M_{ji} < 0$) interactions – see [37] for details. Forty years after May's seminal work [94], Allesina and Tang [37] generalized his result to different interaction topologies. We focus on their results for random (bipartite) mutualistic networks and nested networks, respectively. The comparison between the stability conditions for the two topologies will allow us to gauge the impact of nestedness on the system's local stability.

To generate mutualistic monopartite matrices, for each pair (i, j) of species ($i \neq j$), Allesina and Tang [37] extracted (independently) M_{ij} and M_{ji} from the half-normal distribution $|\mathcal{N}(0, \sigma^2)|$ with probability C . The diagonal elements M_{ii} are always set to $-d$ ($d > 0$) which represents intraspecific competition. Mutualistic bipartite matrices are constructed in a similar way, with the difference that the S species are first split into two equally-sized groups. In bipartite networks, only pairs of species that belong to different groups are connected. The probability to connect a pair of species is given by $\rho = 2C(S - 1)/S$: this choice preserves the expected connectance C of the network. It is essential noticing that both the unipartite and the bipartite mutualistic systems only consider intraspecific competition (M_{ii}) and mutualism ($M_{ij} > 0$ when $i \neq j$), but not interspecific competition.

The stability condition for monopartite mutualistic matrices M is derived based on two observations: (1) the local stability of the matrix is achieved when the largest real part of M 's eigenvalues is negative; (2) the eigenvalue λ_1^M of M with the largest real part is approximately equal to the row-sum $\sum_j M_{ij}$ of the community matrix³⁸

$$\lambda_1^M \simeq \sum_j M_{ij} \simeq -d + (S - 1)C E[|X|]. \quad (96)$$

The stability criterion is, therefore, [37]

$$\frac{(S - 1)C E[|X|]}{\sigma} < \theta. \quad (97)$$

In the numeric example above ($C = 0.1$, $X \sim \mathcal{N}(0, \sigma^2)$, $\theta = 2$), we obtain that the upper bound to the community's size is given by $S < 16$: random mutualistic matrices are less likely to be locally stable than random unstructured matrices. The stability condition for mutualistic bipartite networks is the same as the one for mutualistic monopartite networks.

6.3.3. Stability of nested mutualistic networks

Allesina and Tang [37] also investigated the impact of nestedness on the local stability of a mutualistic community. To generate nested mutualistic matrices M , they fixed the number of links E necessary to achieve the desired connectance $\rho = C$. Then, they ran a filling algorithm that first fills the first row and column of the community matrix, and then arranges the remaining links in such a way that the matrix is perfectly nested. The filled elements M_{ij} of the community matrix are then assigned a value that is randomly extracted from the half-normal distribution $|\mathcal{N}(0, \sigma^2)|$; the diagonal elements are again set to $-d$ ($d > 0$).

Allesina and Tang [37] argued that in nested matrices, stability is undermined by the existence of generalists species that yield larger-than-average row-sum $\sum_j M_{ij}$. Their numerical findings confirmed this intuition: nested bipartite matrices have larger eigenvalues than unstructured bipartite mutualistic matrices (see Fig. 29). In the numeric example above ($C = 0.1$, $X \sim \mathcal{N}(0, \sigma^2)$, $\theta = 2$), the upper bound to the community size is given by $S < 10$: nested mutualistic matrices are less likely to be locally stable than unstructured mutualistic matrices [37].

³⁸ Such result follows from the property that $E[M_{ij}] = C E[|X|]$ for $i \neq j$ and, therefore, $\sum_{j=1}^S M_{ij} \simeq -d + (S - 1)C E[|X|]$. In the matricial notation, the previous equation can be rewritten as $M\mathbf{e} = (-d + (S - 1)C E[|X|])\mathbf{e}$, where \mathbf{e} is the vector whose elements are all equal to one. This means that \mathbf{e} is an eigenvector of M with eigenvalue $-d + (S - 1)C E[|X|]$; one can further prove that this eigenvalue is the one with the largest real part, i.e., $\lambda_1^M \simeq -d + (S - 1)C E[|X|]$. We refer to the Supplementary Material of [37] for more details.

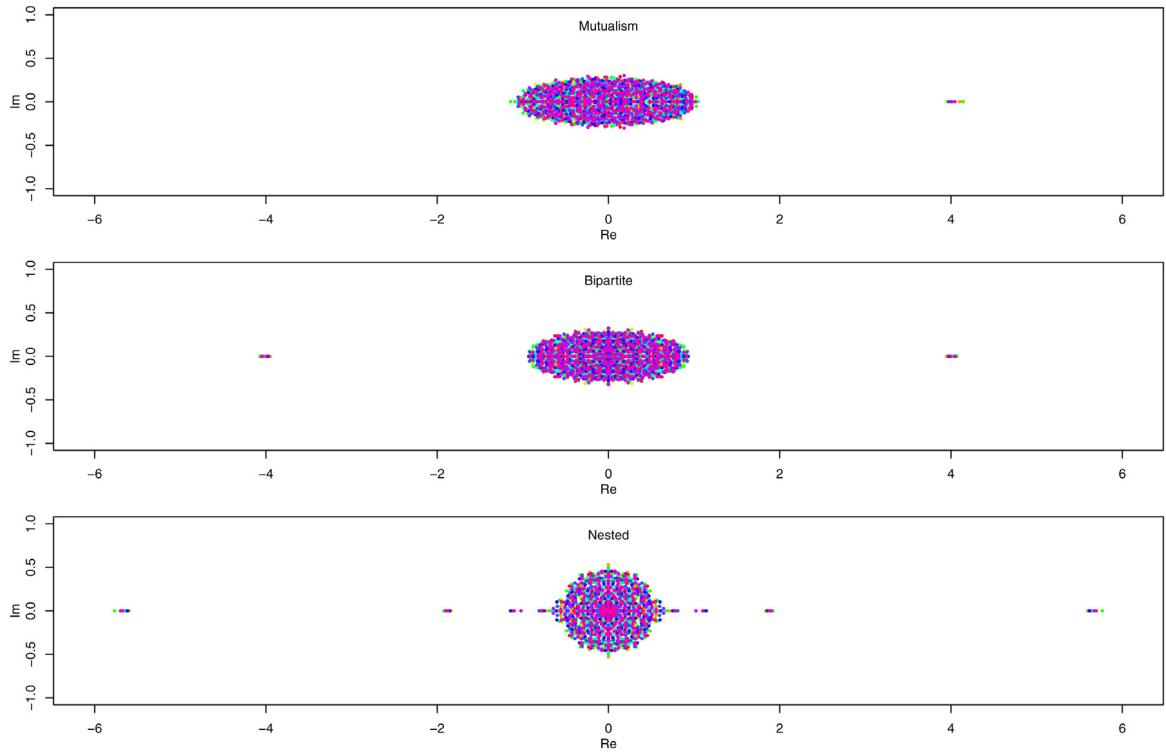


Fig. 29. Local stability in the random-matrix theory approach: real and imaginary parts of the eigenvalues of the community matrix M for an unstructured interacting population (a), a mutualistic bipartite network (b), and a nested network of interactions (c). Nested interactions have eigenvalues with larger real parts than unstructured and mutualistic interactions.

Source: Reprinted from [37].

An analytic explanation of the local instability of nested structures under the random-matrix approach can be obtained by exploiting the connection between perfectly nested structures and the maximal spectral radius of the adjacency matrix [83] (see Section 3.1.4). The community matrices M introduced above can indeed be represented as the sum of an interaction matrix M' with zero diagonal elements, and a diagonal matrix D . For large matrices, the spectral radius of M is given by $\rho(M) = \rho(M') - d$, where d is the average of the elements of D [37]; the system is therefore locally stable only if $\rho(M') < d$. As perfectly nested structures exhibit the largest possible spectral radius (see Section 3.1.4), they are also yield the equilibrium states that are the most unstable against small perturbations [83].

It is important to keep in mind that this result builds on the hypothesis that there are no interspecific competitive interactions. Analytic results on the metapopulation model defined by Eq. (89) also indicate that mutualistic interactions tend to destabilize the system as interspecific competitive interactions vanish [50,340]; however, introducing interspecific competitive interactions leads to a substantially different scenario, where nestedness has a positive impact on biodiversity [50]. The careful reader needs to keep in mind that different assumptions on the nature of interactions lead to different conclusions on the impact of network topology on stability and biodiversity.

6.4. Bottom-line: What is the impact of nestedness on robustness, feasibility, and stability?

To summarize, scholars have investigated the implications of nestedness for various systemic properties related to stability and persistence. Importantly, conclusions obtained by the corresponding studies depend critically on whether interspecific competition is considered, whether feasibility (and not only stability) is ensured in analytical approaches, whether the role of model parameterization is taken into account, and the dimension of stability (linear stability, resilience, local stability) considered. We summarize below some of the main conclusions:

- **Topological robustness** [Section 6.1]. More nested networks are more (less) robust when less (more) connected nodes are more likely to fail and be removed from the system [305]. In bipartite networks, for simple co-extinction cascade processes, rankings of the nodes that produce more nested arrangements of the adjacency matrix better reproduce the structural importance (or vulnerability) of the nodes in both mutualistic networks [71] and country-product networks [72].

- **Feasibility in systems with competition and mutualism** [Section 6.2]. In systems with both competition and mutualism, nestedness enhances the increase of biodiversity due to mutualism [50], and more nested networks are associated with larger feasibility domains, i.e., the network's nodes can co-exist for a larger range of environmental conditions [38,320].
- **Local stability in mutualistic systems without interspecific competition** [Section 6.3]. For purely mutualistic systems (i.e., systems without interspecific competition), under the assumptions of the random-matrix theory approach [37], the maximal size for a nested network such that the system is locally stable is smaller compared to that for a bipartite, unstructured mutualistic network [37]. Even more, when one considers interaction weights, perfectly nested structures (i.e., topologies with the largest spectral radius) are maximally unstable [83].

Beyond the presented results, scholars have also found that more centralized topologies can attenuate the reach of propagating perturbations in the network [377]. A rich scenario emerges where the impact of nestedness depends critically on the systemic property one is interested in. This constellation of potentially contradictory effects of nestedness for different systemic properties (e.g., for systems with competition and mutualism, nestedness tends to favor feasibility but penalize global stability) makes it essential to assess the empirical relevance of different ways of looking at robustness for the actual persistence of real communities. A recent study [320] indicated that at short timescales, empirically observed network dynamics tends to promote feasibility and not stability, yet more research is needed in this direction. We conclude by mentioning that while Saavedra et al. [378] have already extended the structural stability framework to socio-economic networks, this avenue of research remains still largely unexplored.

7. Observing nestedness at the mesoscopic scale

In the previous Sections, we have studied nestedness at a *macroscopic* level by defining and measuring it as a property related to the connectivity patterns of *all* nodes (see Fig. 30C–G). Besides, we have pointed out that ecological and socio-economic networks can exhibit other interesting structural patterns. Some of them (like heterogeneous degree distribution, disassortativity, core–periphery structure) are often observed together with nestedness (see Section 4.1), whereas others – like gradient structure (Fig. 30A–E), modular or compartmental structure (Fig. 30 B–F) – seem to be incompatible with nestedness.

On the other hand, real ecological processes that lead to the observed structural patterns are typically confined by inherent boundaries due to diverse reasons [96]. For example, for ecological datasets that include observations made in distinct spatial regions, it is interesting to assess not only the degree of nestedness of the complete network, but also the degree of nestedness of subgraphs that correspond to the different geographical regions [42,379]. Indeed, variations in nestedness across different geographical locations might reflect different geographical histories of the subgroups of species that inhabit them [379].

Boundaries and constraints play a significant role in socio-economic systems as well. For instance, the probability that two countries have a trade relationship decreases with their geographical distance [381], and geographical effects may affect the compartmentalized structure of trade [104] and communication networks [382]. Both in ecological and socio-economic systems, if the mechanisms that drive the internal dynamics inside network compartments is ruled by one of the processes that lead to nested structures (see Section 5), one might observe as a result “combined” compartmentalized structures where each compartment exhibits, internally, a nested structure (see Fig. 30D–H).

Lewinshon et al. [96] already recognized the potential importance of the combined structure depicted in Fig. 30D–H, emphasizing that processes that generate ecological networks typically “operate within a framework of boundaries set by morphological, functional or phylogenetic constraints”, which implies that “most species, but not necessarily all, will preferentially establish links in a given compartment and, within that compartment, their host range or fauna will be conditioned by more proximate factors” [96]. Lewinshon et al. [96] applied correspondence analysis techniques to discriminate between the different structural patterns showed in Fig. 30 – we refer to [96] for details. Subsequent studies used standard nestedness and community-detection tools [223] to measure the level of nestedness for different geographical subregions [379], to assess the overall correlation between nestedness and modularity [39], and to measure the degree of nestedness of the modules detected by modularity-maximization algorithms [380,383].

Here, we refer to *nestedness at the mesoscopic scale* whenever nestedness is studied not as a property of the whole network (macroscopic scale), but of one or more than one of its subgraphs. Methodologies specifically aimed at detecting nested patterns at the mesoscopic scale have received growing interest over the last five years [40,42,43,384,385], in relation to three main problems: (1) quantifying the level of nestedness of a given subgraph [42] (see Section 7.1); (2) detecting, for a given network, the largest subgraph that exhibits an internally nested topology [43,384] (see Section 7.2); (2) partitioning the network into a set of compartments, such that each compartment exhibits, internally, a nested [40] or a core–periphery-like [44,385] topology of interactions (see Section 7.3).

In this Section, we review the recent progress that has been made in relation to these three problems. In spite of the massive amount of works on nestedness, our understanding of nestedness at the mesoscopic scale is still at its infancy. Nevertheless, the obtained results in this direction suggest that future research might shift the focus from detecting system-wide, “macroscopic” nestedness to detecting confined, “mesoscopic” nested structures. Detecting these structures, understanding the mechanisms behind their emergence and their implications may lead to new insights on the organizing principles of ecological and socio-economic networks.

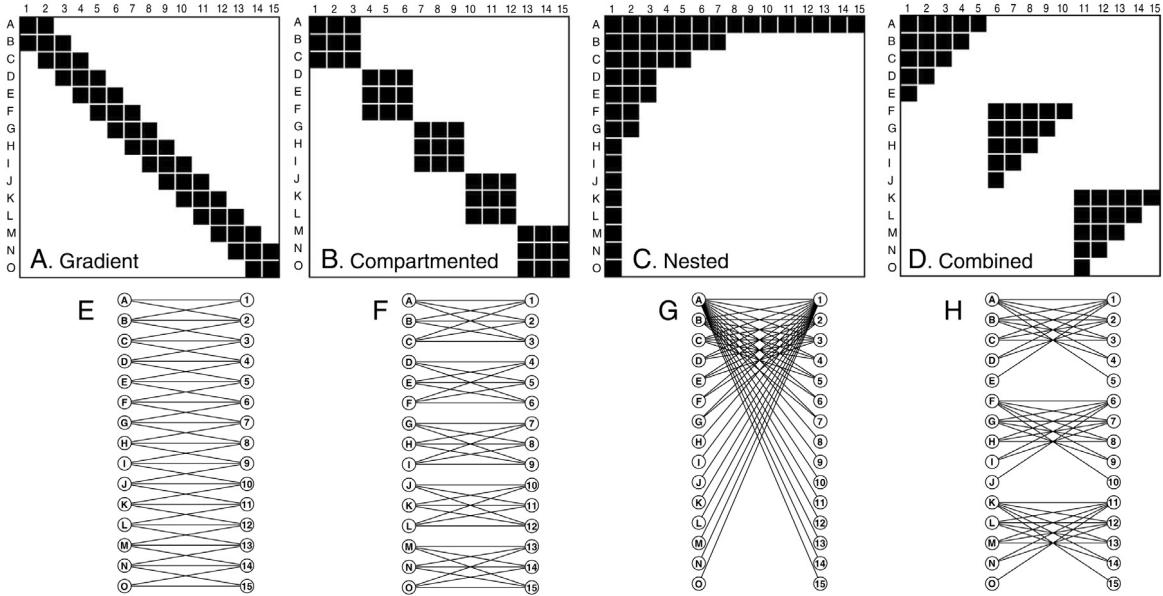


Fig. 30. Illustration of different topological patterns. Panels A–E illustrate a gradient structure where each node only interacts with adjacent nodes in the matrix. While these patterns have been investigated in vegetation ecology [96], they are rarely found in plant–animal interaction networks and socio-economic networks. Panels B–F illustrate a “compartmented” [245,246] (or “modular” [222]) structure where the network can be partitioned into well-defined “compartments”, or “modules”, or “blocks”, or “communities”. There exist many strategies to detect this kind of structures [222,223]; extant detection techniques usually do not take into account the possible internal topology of the compartments. Panels C–G illustrate a (perfectly) nested structure, which is the main focus of this review and is defined in the Introduction section. Panels D–H illustrate a “combined” (or “nested–modular” [380], or “in-block nested” [40]) structure where the network is compartmented, but each compartment (or “block”) exhibits an internally nested structure.

Source: Reprinted from [96].

7.1. Quantifying the degree of nestedness of a subgraph

Given two subgraphs $S_1 \subset \mathcal{G}$ and $S_2 \subset \mathcal{G}$ of a network \mathcal{G} of interest, one might be tempted to assess their relative level of statistical significance by simply comparing the z-scores of their degree of nestedness (as determined by a suitable nestedness metric, like NODF or matrix temperature). However, in principle, the resulting z-scores may depend on matrix size. To overcome this obstacle, Strona et al. [42] developed a procedure to compare the levels of nestedness of subgraphs of different size, and subsequently applied it to detect structural differences between Anatolian and Cyclades islands in the Aegean islands.

Suppose that we are interested in assessing the statistical significance of nestedness for a given subgraph $S \subset \mathcal{G}$. To fix ideas, we consider NODF as our nestedness metric, yet in principle, the steps below apply to any nestedness metric. The procedure by [42] requires five steps (see Fig. 1 in [42] for a diagrammatic representation of the procedure): (1) We calculate the overall level of nestedness of the network \mathcal{G} , i.e., the z-score of NODF (based on the expected value and the standard deviation of NODF in an ensemble of suitably randomized networks); (2) We randomly partition the network into matrices of random size, and we calculate the z-score of NODF for each of them (in the same way as in (1)); (3) We regress the obtained subgraphs’ z-score against subgraph size; (4) The obtained regression coefficients determine, for each subgraph size, the expected NODF for subgraphs of that size; (5) The observed z-scores for any subnetwork can be compared with the random expectation for subgraphs of the same size. We refer to [42] for additional details on these five steps.

The procedure outlined above can be used to detect differences between different subgraphs of ecological interest. Strona et al. [42] applied it to species distribution patterns in the Aegean islands, a group of islands in the Aegean sea. Interestingly, based on ecological considerations [379], the Aegean islands can be divided into two groups: Anatolian and Cyclades. The procedure by Strona et al. [42] revealed that the Cyclades and the Anatolian islands tend to be substantially more and less nested, respectively, than expected for a randomly-extracted subgraph of equal size. Importantly, as different processes can lead to different degrees of nestedness in the resulting species distribution patterns, variation in nestedness across subgraphs indicate that different parts of the systems have undergone different ecological processes [42,379].

7.2. Finding the largest nested subgraph

Besides evaluating the degree of nestedness of specific subgraphs, one might be interested in finding the largest subgraph (referred to as *nested component* [43,384]) that exhibits a significantly nested structure. The relevance of this problem stems from the fact that, with increased data availability or longer observational periods, not necessarily all nodes in the dataset considered are to be part of a nested arrangement. Furthermore, traditional approaches to nestedness detection are top-down: It is up to the researchers to determine the realm of observation, in some cases in an arbitrary manner. Instead, it might be more appropriate to include all the available data into the analysis, and determine algorithmically the largest subset of the data that exhibit a nested structure. This problem has been recently tackled by Grimm and Tessone [43,384] who introduced an algorithm called NESTLON (NESTedness detection based on LOcal Neighborhood). Below, we describe the algorithm for unipartite networks, yet the algorithm can be applied to bipartite networks as well.

Essentially, NESTLON sequentially “constructs” the nested component $\Gamma_{\text{nest}} \subseteq \mathcal{G}$. The nested component is initialized with an empty set of nodes. We start from the node i_1 with the largest degree³⁹; this node is a candidate to belong to the nested component. The main idea of NESTLON is to assess whether the neighborhood of this node includes the neighborhood of lower-degree nodes (up to a “confirmation ratio” θ_{nest} which allows for a tunable level of deviation from a perfectly nested structure). If this is true for a sufficiently large fraction of i_1 ’s neighbors (where such minimum fraction is a parameter of the algorithm), then the neighborhoods of i_1 ’s neighbors are nested inside i_1 ’s neighborhood (up to an acceptable level), and i_1 is added to the nested component, which now includes one node. i_1 ’s neighbors are now candidates for the same evaluation: one considers their neighbors, and assess whether, for a sufficient fraction of them, their neighborhoods are sufficiently nested inside i_1 ’s neighborhood. If the neighborhoods of i_1 ’s neighbors are not sufficiently nested inside i_1 ’s neighborhood, then we do not add i_1 to the nested component, and we evaluate the neighborhood of the neighbors of the node i_2 with the second largest degree. The procedure continues iteratively until it is not possible to add more nodes to the nested component.

Clearly, in a perfectly nested network, $\Gamma_{\text{nest}} = \mathcal{G}$. For networks that are not perfectly nested, NESTLON provides us not only with the nested component, i.e., with the largest subgraph that exhibits a nested structure, but also a new metric for nestedness: the relative size $\mu_{\text{nest}} = |\Gamma_{\text{nest}}|/N$ of the nested component. Indeed, $\mu_{\text{nest}} = 1$ corresponds to a perfectly nested network, whereas $\mu_{\text{nest}} = 0$ indicates a maximally non-nested network where no pair of nodes respects the nestedness condition. Results on synthetic networks suggest that this metric may be better suited than NODF and BINMATNEST for nestedness quantification in sparse and high-density matrices. It is worth noticing that the algorithm operates based on local information, which makes the algorithm scalable with system size.

7.3. Detecting in-block nestedness

Think to the communication patterns of the students in a given school. Likely, the students of a given class tend to communicate more often with the students of the same class than with the students of a different class. Let us assume that the communication patterns within a class are nested — i.e., the friends of a given student i tend to form subsets of the friends of students who have more friends than i . This pattern of interactions can be described as an *in-block nested structure*: differently from nested and modular structures, in-block nested structures are characterized by compartments of nodes that *internally* exhibit a nested pattern of interactions (Fig. 30D–H).

In relation to the previous example, we have no a priori guarantee that the communication patterns within the individual classes are nested. A natural question emerges: how to properly quantify the level of in-block nestedness in a given network? To address this question, we need to simultaneously assess two properties: (1) How well the network can be partitioned into compartments; (2) How nested is the communities’ internal topology.

A simple way to address both points would be to first identify the network’s compartments, or communities, through a modularity-optimization algorithm, and then measure the level of nestedness of the detected communities (Section 7.3.1). However, this approach can fail as modularity-optimization algorithms may be unable to detect blocks that exhibit a significantly nested structure (Section 7.3.1). To overcome this limitation, we need methods that specifically aim to detect in-block nested structures; such methods are the main topics of Sections 7.3.2 and 7.3.3.

7.3.1. Nestedness of network modules

A tempting procedure to evaluate the presence of communities with internally nested structure is to first identify the communities via some well-established community-detection algorithm [222,223], and then measure the level of nestedness within each detected community. Flores et al. [383] evaluated a large phage–bacteria interaction network composed of 215 phage types with 286 host types sampled from geographically separated sites in the Atlantic Ocean. They found that this interaction network is highly modular, and evaluated the level of (internal) nestedness of the individual modules. They found that some of the modules are significantly nested, whereas others are not (Fig. 31). They observed [383] that even in situations where nestedness is unlikely at the global scale, one may still observe nestedness “at

³⁹ To simplify the discussion, let us assume that the node with degree equal to the largest degree k_{\max} in the network is unique. The extension to networks where more than one node have a degree equal to k_{\max} is straightforward [43].

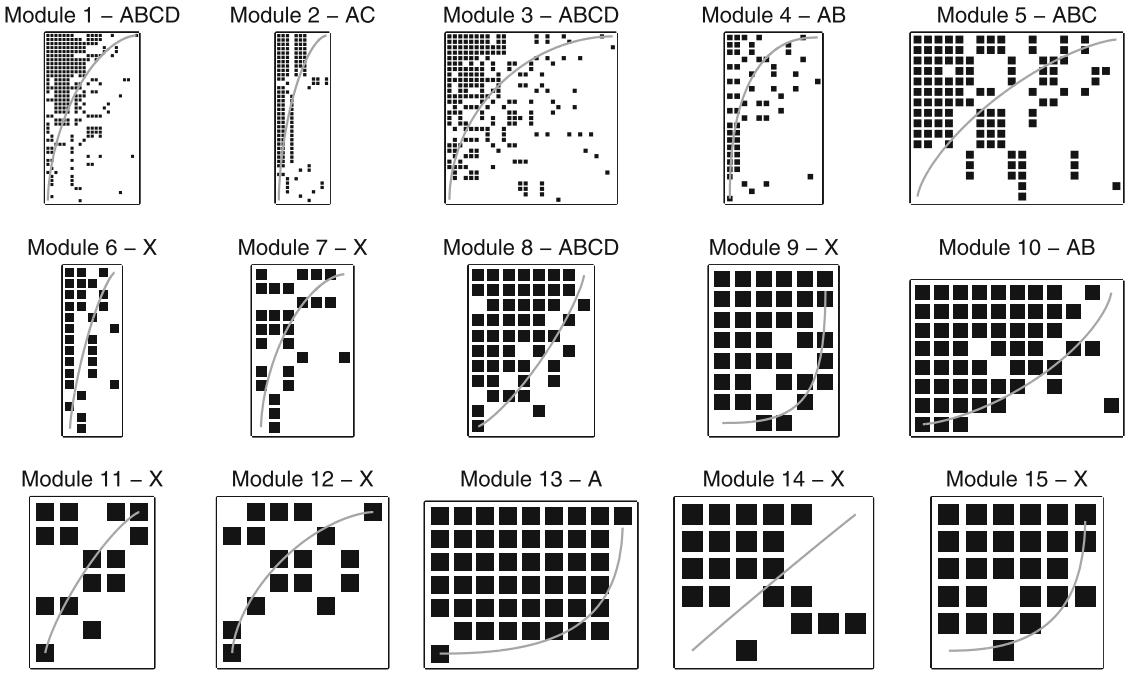


Fig. 31. Internal structure of the modules detected through modularity optimization in the phage–bacteria infection network analyzed by Flores et al. [383]. The gray line represents the isocline of perfect nestedness by the NTC algorithm (see Section 3.1.2). Each module is marked with labels that denote the significance of its internal nested structure: A and B mean that the module is nested according to the NTC under the EE and PP null model, respectively (see Section 3.2.1); C and D mean that the module is nested according to the NTC under the EE and PP null model, respectively; X means that no significance was found.

Source: Reprinted from [383].

the local scale” (i.e., within modules), and they pointed out the need to develop metrics that can disentangle nestedness and modularity in networks. Beckett and Williams [380] referred to the presence of nestedness inside network modules as a “nested–modular” structure, and proposed a co-evolutionary model based on genetic matching to explain its emergence.

On the other hand, Solé-Ribalta et al. [40] generated synthetic data with planted in-block nested structures (see Fig. 1 in [40]), and found that even in structures with no inter-community links, modularity maximization can fail to uncover the ground-truth communities. This happens because, in sparse nested structures, modularity maximization tends to cluster together the high-degree nodes due to the high density of links among them. This result suggests that while sequentially applying two well-established community and nestedness detection techniques is appealing, such approach is inappropriate when we wish to detect communities that exhibit an internally nested structure, as modularity-optimization algorithms might incapable of detecting the ground-truth communities. This finding calls for methods that take into account simultaneously the community structure of the network and the internal nested topology of the communities.

7.3.2. Detecting in-block nested structures

To overcome the limitation of the procedure described in Section 7.3.1, Solé-Ribalta et al. [40] introduced a quality function specifically aimed to detect in-block nested structures which takes into account the fact that the nodes can be partitioned into blocks, and each block exhibits an internally nested structure. In bipartite networks, the resulting quality function, called *in-block nestedness* [40], is defined as

$$\mathcal{I} = \frac{2}{N+M} \left\{ \sum_{ij} \frac{O_{ij} - \langle O_{ij} \rangle}{k_j(|E_i| - 1)} \Theta(k_i - k_j) \delta(E_i, E_j) + \sum_{\alpha\beta} \frac{O_{\alpha\beta} - \langle O_{\alpha\beta} \rangle}{k_\beta(|E_\alpha| - 1)} \Theta(k_\alpha - k_\beta) \delta(E_\alpha, E_\beta) \right\}. \quad (98)$$

In unipartite networks, the previous expression reduces to [386]

$$\mathcal{I} = \frac{2}{N} \sum_{ij} \frac{O_{ij} - \langle O_{ij} \rangle}{k_j(|E_i| - 1)} \Theta(k_i - k_j) \delta(E_i, E_j), \quad (99)$$

The \mathcal{I} quality function can be interpreted as a generalization of the normalized NODF function defined by Eq. (25). The contribution to \mathcal{I} from a given pair (i, j) of nodes has indeed the form $\Theta(k_i - k_j) O_{ij}/k_j$; the additional elements with respect to the NODF function are (1) the $\delta(E_i, E_j)$ term which restricts the contributing pairs of nodes to those whose nodes belong

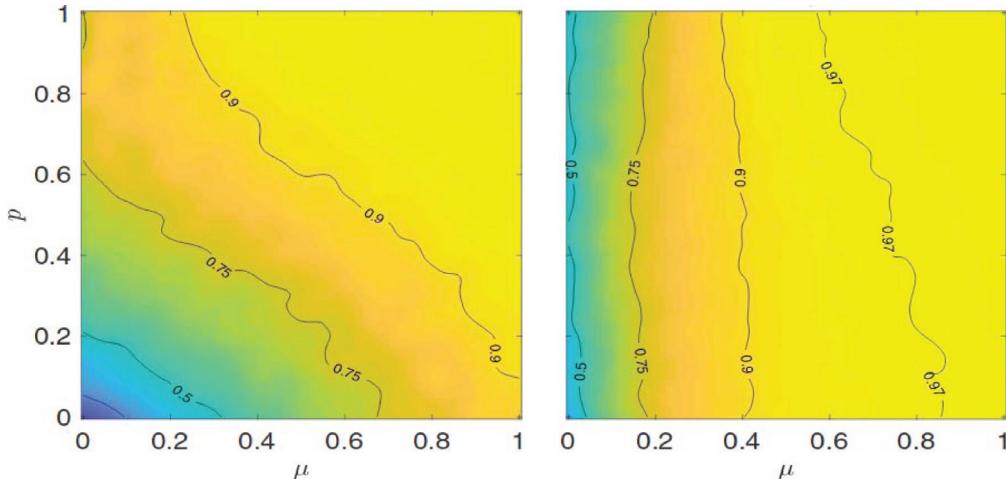


Fig. 32. An illustration of the ability of in-block nestedness optimization (left panel) and modularity optimization (right panel) to reconstruct planted compartments in networks with tunable fraction of inter-compartment links (parameter μ) and level of nestedness within the compartments (parameter p – the smaller, the more internally nested the compartments). The reconstruction ability of the two methods is quantified by the Normalized Variation of Information (NVI) [222] between the partition detected by the method and the ground-truth partition of the nodes. In-block nestedness provides a significantly better reconstruction in the low- p region.

Source: Adapted from [40].

to the same block; (2) the normalization factor $|\mathcal{E}_i| - 1$ (where $|\mathcal{E}_i|$ denotes the number of nodes that belong to the block \mathcal{E}_i to which node i belongs) which ensures that the in-block nestedness \mathcal{I} of a maximally nested structure tends to one in the thermodynamic limit [386]. The maximization of \mathcal{I} (performed through a biologically inspired optimization algorithm in [40]) provides us with both the optimal partition Ξ of the nodes into blocks, and the (optimal) degree of in-block nestedness \mathcal{I} of the network. A large value of \mathcal{I} indicates that it is possible to partition the network into compartments, each of them with an highly-nested internal structure.

Solé-Ribalta et al. [40] validated the \mathcal{I} function by applying it to synthetic networks with a planted in-block nested structure. The model that generates their benchmark graphs features two parameters: the level p of nestedness within each block ($p = 0$ and $p = 1$ correspond to a perfectly nested and a random structure, respectively); the fraction μ of inter-block links (the correct reconstruction of the blocks is increasingly difficult as μ increases). One would expect a good method for the detection of IBN structures to correctly reconstruct the planted IBN blocks in the low- p and low- μ parameter region.

They found (Fig. 32) that modularity-optimization fails to correctly reconstruct the planted IBN blocks in a large portion of the parameter space. For relatively sparse networks, modularity is confused even by a tiny amount of inter-block links. By contrast, optimization of the \mathcal{I} quality function leads to an accurate reconstruction of the planted IBN blocks for a broad portion of the low- p and low- μ region of the parameter space. As μ and p gradually increase, the performance of the \mathcal{I} -optimization algorithm gradually deteriorates, yet in-block nestedness maximization substantially outperforms modularity maximization for networks generated with low values of p .

In real data, they found (Fig. 33) that nestedness and in-block nestedness are somewhat independent properties. The correlation between nestedness and modularity is indeed small for both unipartite and bipartite networks, and there is no specific range of nestedness values where in-block nestedness tends to be large. Besides, the modularity's resolution limit impairs its ability to identify modules that are smaller than a certain scale which depends on network properties [241]. A numerical investigation [40] suggests that differently from modularity, the in-block nestedness function might not exhibit a similar resolution limit.

7.3.3. Detecting multiple core-periphery pairs

In the previous Section, we have explored particular network structural patterns where the network can be partitioned into blocks of nodes, and the connectivity patterns within each block exhibit a nested organization. We have seen that the maximization of a suitable quality function, called in-block nestedness [40], allows us to quantify the significance of this pattern. Motivated by this observation, we can ask ourselves: can we detect other mesoscopic structural patterns with a similar approach?

In this Section, we consider the detection of multiple core-periphery pairs [385]. The topic is relevant to the present review as a core-periphery structure can be considered as a special case of a nested structure (see Section 4.1.3). As we have seen in Section 4.1.3, a core-periphery structure is composed of a “core” of nodes that are densely connected with all the other nodes, and a “periphery” of nodes that tend to be only connected with the nodes in the core. Kojaku and Masuda [130] generalized the notion of core-periphery by building a quality function that considers the possibility of multiple core-periphery structures within the same network.

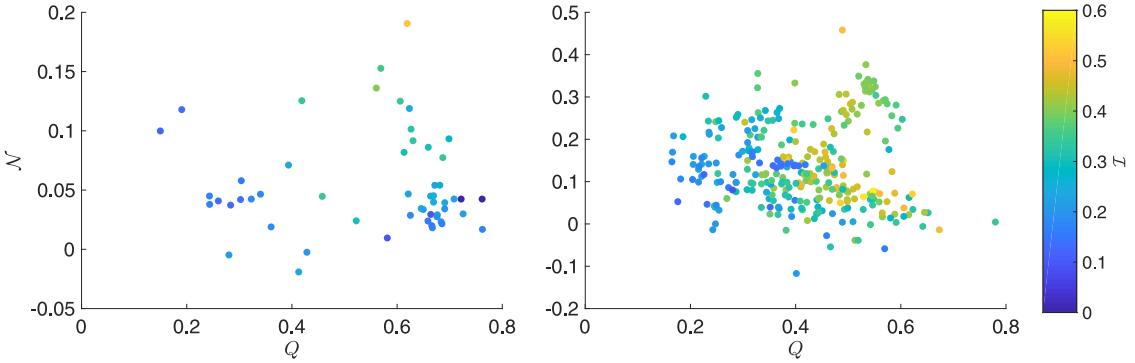


Fig. 33. The relation between nestedness \mathcal{N} , modularity Q , and in-block nestedness \mathcal{I} in unipartite and bipartite empirical ecological, social or urban networks. Each dot represents a network, and its color represents the level of in-block nestedness \mathcal{I} . There is no evident pattern in the mutual dependence of these three properties.

Source: Reprinted from [40].

Quality function for multiple core–periphery structures. Kojaku and Masuda [385] introduced a structural pattern where each node belongs to a given block, and each block *internally* exhibits a core–periphery structure. In a similar way to Borgatti and Everett (see Section 4.1.3), they defined an ideal *multiple* core–periphery structure A^{MCP} where the network is partitioned into disjoint blocks $\Xi \in \{1, \dots, B\}$, and each blocks exhibits an ideal core–periphery structure. Each node i unambiguously belongs to one of the blocks, which is therefore denoted as Ξ_i . In a similar way to Section 4.1.3, for each node i , one introduces a binary variable x_i such that $x_i = 1$ and $x_i = 0$ for core and peripheral nodes, respectively – the core/peripheral role is now played by each node *within* the block to which the node belongs.

In terms of the Ξ and \mathbf{x} variables, $A_{ij}^{MCP} = (x_i + x_j - x_i x_j) \delta(\Xi_i, \Xi_j)$. The novel element with respect to A_{ij}^{CP} (defined for a single, macroscopic core–periphery structure – see Section 4.1.3) is the Kronecker delta $\delta(\Xi_i, \Xi_j)$ that only allows nodes of the same block to be connected with each other. The expected number of links that are present in both the original network A and the ideal multiple core–periphery network A^{MCP} is simply $\sum_{(i,j)} A_{ij}^{MCP}$. By comparing this number with its expected value $p \sum_{(i,j)} A_{ij}^{MCP}$ for an Erdős–Rényi network, we obtain the $Q^{MCP;ER}$ quality function [385]

$$Q^{MCP;ER}(\Xi, \mathbf{x}) = \sum_{(i,j)} (A_{ij} - p) A_{ij}^{MCP}(\Xi, \mathbf{x}) = \sum_{(i,j)} (A_{ij} - p)(x_i + x_j - x_i x_j) \delta(\Xi_i, \Xi_j). \quad (100)$$

By comparing $\sum_{(i,j)} A_{ij}^{MCP}$ its expected value $p \sum_{(i,j)} A_{ij}^{MCP}$ for a random network with the same degree sequence as the original one (i.e., with the expected value under the configuration model), we obtain the $Q^{MCP;CM}$ quality function [44]

$$Q^{MCP;CM}(\Xi, \mathbf{x}) = \frac{1}{2E} \sum_{(i,j)} \left(A_{ij} - \frac{k_i k_j}{2E} \right) (x_i + x_j - x_i x_j) \delta(\Xi_i, \Xi_j). \quad (101)$$

The maximization.⁴⁰ of $Q^{MCP;ER}$ and $Q^{MCP;CM}$ (performed through a label switching heuristic [44,385]) provides us with: (1) the set of optimal Ξ values – i.e., the partition of the nodes into blocks; (2) the set of optimal \mathbf{x} values – i.e., the detected within-block cores. A large value of $Q^{MCP;ER}$ or $Q^{MCP;CM}$ indicates a structure that has large overlap with an ideal multiple core–periphery structure.

Interestingly, the $Q^{MCP;CM}$ reduces to the modularity function (defined by Eq. (54)) when all nodes are core nodes. One can also define the quality of a detected block Ξ as [44]

$$q(\Xi) = \frac{1}{2E} \sum_{(i,j)} \left(A_{ij} - \frac{k_i k_j}{2E} \right) (x_i + x_j - x_i x_j) \delta(\Xi_i, \Xi) \delta(\Xi_j, \Xi). \quad (102)$$

A block Ξ is deemed as statistically significant if its quality $q(\Xi)$ substantially exceeds the expected value for blocks of the same size $|\Xi|$ in networks randomized with the configuration model. In a similar way to the modularity function [241], the $Q^{MCP;CM}$ function suffers from a resolution limit that impairs its ability to detect core–periphery pairs below a certain size [44]. Such a limitation can be overcome by introducing a multi-resolution quality function where the expected overlap between the observed and ideal multiple-core–periphery network is modulated by a resolution parameter γ – see [387] for details.

To validate the $Q^{MCP;ER}$ function, Kojaku and Masuda [44] compared its detected core–periphery structure with that from three alternative methods:

⁴⁰ The corresponding code is available at: https://github.com/skojaku/km_config/.

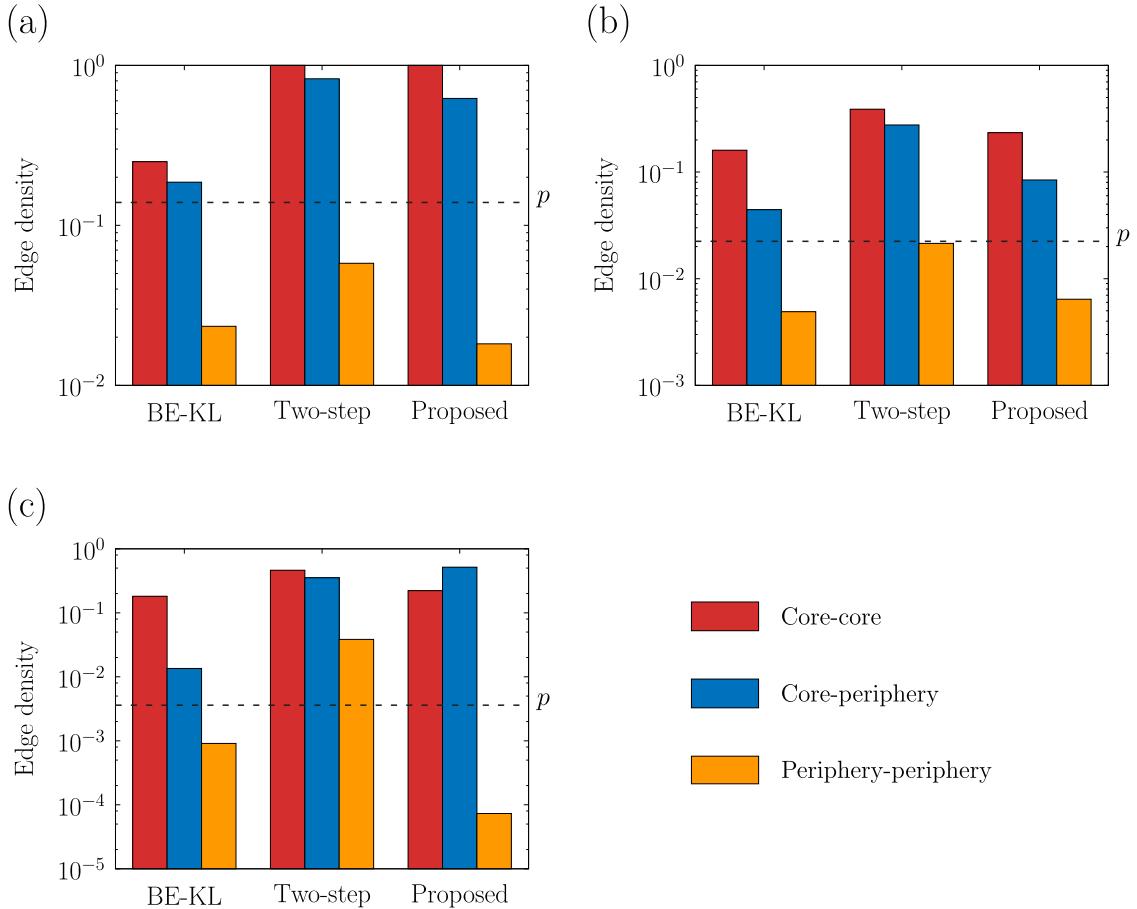


Fig. 34. A comparison of core–periphery detection methods in three empirical networks: (a) Zachary’s karate club; (b) the network of political blogs on the 2004 US presidential elections; (c) the World air transportation network. The methods are compared with respect to the density of links between different classes (core/periphery) of nodes, where the two classes are determined by the three different methods described in the main text: BE-KL, Two-step, and Q^{MCP} -maximization method (referred to as “proposed” in the figure). Compared to BE-KL and Two-step, the Q^{MCP} -maximization produces core–periphery pairs that exhibit substantially lower densities of periphery–periphery edges, in agreement with the notion of core–periphery.

Source: Reprinted from [385].

- The original Borgatti–Everett (BE) core detection algorithm which aims to detect one single core–periphery structure. As they maximized the corresponding quality function Q^{CP} (defined by Eq. (47)) through the Kernighan–Lin algorithm [388], they referred to the method as the BE-KL method.
- A “two-step algorithm” where one first partitions the nodes through modularity optimization (performed with the Louvain algorithm [239]) and, at the same time, applies the BE core detection algorithm to the whole network. The core nodes detected by the BE algorithm are then assumed to be the core nodes within each module.
- A “divisive algorithm” where one first partitions the nodes through modularity optimization (performed with the Louvain algorithm [239]), and then applies the BE core detection algorithm to each of the detected modules, in a similar spirit to the procedures described in Section 7.3.1.

Results on three empirical networks (see Fig. 34) show that with respect to these methods, the Q^{MCP} maximization is the only one that produces core–periphery blocks such that – consistently with the main idea behind core–periphery structures – the density of intra-periphery edges is significantly smaller than the expected density for a random graph. Besides, the Q^{MCP} maximization was found to be the only algorithm that reveals at the same time: (1) Groups of nodes that share the same property, as revealed by comparing the obtained node partitions with metadata; (2) Nodes’ role within each group, as revealed by direct inspections of the core and the peripheral nodes within each detected block. Recently, the $Q^{MCP;ER}$ maximization has been applied to detect structural changes in the interbank market [130].

Results on synthetic networks [44] indicate that the $Q^{MCP;CM}$ -maximization algorithm outperforms both the divisive algorithm and the $Q^{MCP;ER}$ -maximization algorithm in reconstructing planted core–periphery pairs. In particular, differently from the divisive and the $Q^{MCP;ER}$ -maximization algorithm, the $Q^{MCP;CM}$ -maximization algorithm is able to reconstruct

ground-truth blocks even when a substantial fraction of core nodes are not hubs and a substantial fraction of peripheral nodes are hubs [44].

8. Outlook

This review has focused on a network structural property that has been introduced in ecological studies of species-site spatial networks [21] and has subsequently attracted great interest from diverse research fields: nestedness. It has examined existing studies on nestedness by following three steps: (1) introducing the main methodologies to detect nestedness (*observation*), (2) presenting the main mechanisms proposed to explain the pervasiveness of nestedness in ecological and economic systems (*emergence*), and (3) assessing the main implications of nestedness for network topological robustness, feasibility, and stability (*implications*). In principle, this three-step framework provides us with a conceptual roadmap that can be equally adopted to survey existing studies on other structural patterns of real-world networks (e.g., modularity [223] and rich-club structure [389]).

Stimulated by our review, a natural question emerges: which are the next steps for research on nestedness and, more in general, on the architecture of ecological and socio-economic systems? As we have often recognized throughout this review, some questions related to the structure of ecological and socio-economic networks have not yet found a universal answer. Besides methodological questions (e.g., how to best measure the level of nestedness of a given network, and its statistical significance?), we argue that questions related to the emergence and implications of nestedness and other structural patterns of ecological and socio-economic systems need additional efforts in order to be answered.

As for the emergence of nestedness, we already explored (Section 5) various mechanisms that can, in theory, drive a network from a random to a nested topology of interaction, or growth mechanisms that can assemble nested networks. However, studies that adopt statistical techniques to validate such mechanisms on time-stamped data are non-existent. As inferring the dynamical generative process from the structure of the final network is not possible, we point out a thorough temporal analysis of the dynamics of ecological and socio-economic networks as the only way to single out the mechanisms that are responsible for the empirically observed nestedness, and rule out those that do not take place. Toward this direction, experimental approaches [249] might play an increasingly central role.

It is essential to notice that our review has focused on the nestedness of unipartite and bipartite networks. The extension of nestedness to temporal and multilayer networks is still largely unexplored in the literature, in contrast with other structural properties that have been already generalized (e.g., centrality metrics [291,390] and modularity [243]). While preliminary works in this direction can be already found [108,391], we are still far from having a complete understanding of the causes and implications of nestedness in multilayer networks. This is an important research gap because the fundamental role by different types of interactions is increasingly recognized in ecology [392–394] and trade analysis [108,395]. We predict that several future studies will attempt to understand the role of nestedness (and related structural patterns) in multilayer and temporal networks.

Finally, while nestedness has traditionally been considered as a macroscopic network property that involves all the nodes of a given network (Section 3), a number of recent techniques have been developed to detect nestedness at a mesoscopic scale instead of the global one (Section 7). Nevertheless, our knowledge of the performance of these detection methods, and our understanding of the causes and implications of such mesoscopic structures are still at their infancy. This class of methods might play an increasingly central role in nestedness analysis as real systems are subject to various types of constraints [96] which makes it natural to look for patterns of interaction organization (like nestedness) at a “local” scale and not at the global one.

We conclude by stressing that our review is far from being exhaustive. Besides introducing methodologies for the detection of nestedness, it mostly aimed to introduce the essential “physical” aspects of nestedness: stochastic mechanisms that lead to its emergence, and its impact on dynamical processes on the network. We hope that this review will be a useful compendium not only for physicists, but also for ecologists and economists. In particular, ecologists (economists) can gain inspiration from mechanisms and techniques introduced in economics (ecology) and physics, which might lead to unexpected applications of nestedness and network analysis which transcend traditional disciplinary boundaries.

Acknowledgments

We wish to thank Javier Borge-Holthoefer, Matúš Medo, Albert Solé-Ribalta for the many inspiring discussions on nestedness and the detection of structural patterns in networks. We are grateful to Orazio Angelini, Tomaso Aste, Clàudia Payrató Borràs, Luciano Pietronero, Andrea Zaccaria for their feedback on the first draft of this article and their invaluable suggestions which helped us to improve several sections of the article. This work has been supported by the Science Strength Promotion Program of UESTC and by the URPP Social Networks. MSM and CJT acknowledge financial support from the Swiss National Science Foundation (Grant No. 200021-182659). MSM acknowledges financial support from the UESTC professor research start-up (Grant No. ZYGX2018KYQD21). ZMR was partially supported by the National Natural Science Foundation of China (Grant No. 61803137), the Hangzhou Normal University Research Funding Project, the Qiantang River Talents Plan (Grant No. QJD1803005), and the Foundation of High level overseas returnees (team) in Hangzhou for Pioneering Innovation Program. JB acknowledges financial support from the Swiss National Science Foundation (Grant No. 31003A-169671).

Appendix A. Software and relevant datasets

We stress that the URLs provided below were active at the time when this article was published. There is no guarantee that these URLs will be still active at the time when this article will be read.

A.1. Software for nestedness and network analysis

Standard, general packages for network analysis are: `NetworkX`, `igraph`, `graph_tool`, among many others. The R package `bipartite` was specifically designed for the analysis of bipartite networks, and it implements various nestedness metrics. Besides, the R package `vegan` was designed for the analysis of ecological communities, and it also incorporates various nestedness metrics.

A.1.1. Metrics and null models

Besides general packages for network analysis, due to the massive popularity of nestedness analysis in ecology, scholars have developed specific packages to compute metrics for nestedness together with their statistical significance. Among them, we mention `FALCON` [<https://github.com/sjbeckett/FALCON>] and `Nestedness for Dummies` [<http://ecosoft.alwaysdata.net/download/>]. Both packages implement various nestedness metrics, null models, and significance tests. We refer the interested reader to [160,396] for more information.

A.1.2. Nested network generation

When we wish to generate perfectly nested or highly-nested networks, we have two options: (1) decide the shape of the separatrix in the adjacency/incidence matrix (e.g., Fig. 1), and add links above the chosen line; (2) adopt one of the generative mechanisms described in Section 5. As for (2), a function to generate threshold graphs (i.e., based on the discussion in Section 2.1, perfectly nested networks) based on the mechanisms described in Section 5.1 can be found in the R package `netrankr` [`threshold_graph` function].

A.1.3. Community detection and nestedness at the mesoscopic scale

The main network-analysis packages implement several algorithms for the detection of communities. Among the modularity maximization algorithms, the `Combo` algorithm turned out to be more effective than other existing methods [240]; the respective code can be found at http://senseable.mit.edu/community_detection/. For a recent and exhaustive list of available software for community detection, we refer the reader to Section 5 of the review article [223]. Importantly to nestedness analysis, recent works have started investigating nestedness and core-periphery as a mesoscopic property of networks, as extensively discussed in Section 7. The code that implements the maximization of the multiple core-periphery function by Kojaku and Masuda [44] (see Section 7.3.3) can be found at https://github.com/skojaku/km_config/.

A.2. Datasets of socio-economic and ecological networks

A.2.1. World trade data

World Trade datasets can be downloaded from several sources. The COMTRADE dataset [available at <https://comtrade.un.org/>] is curated by the United Nations Statistics Division, and it includes bilateral trade flows declarations for several decades [45]. World Trade datasets from various sources are also available at the Observatory of Economic Complexity collection [<https://atlas.media.mit.edu/en/resources/data/>] and the Economics Web Institute webpage [<http://www.economicswebinstitute.org/worldtrade.htm>]. It is essential to note that such datasets are usually noisy. For example, in the COMTRADE data, the declared yearly import volume for a given exchange often does not match the declared export volume, which makes it necessary to develop data sanitation procedure. Discussing the topic in detail goes beyond the scope of this review, but we refer the interested reader to [45,113] for relevant discussions.

A.2.2. Ecological spatial networks

The original Nestedness Temperature Calculator (NTC) software by Atmar and Patterson [22] can be downloaded at the link [http://priede.bf.lu.lv/ftp/pub/TIS/datu_analiiize/Nestedness/about.html]. The program includes 294 presence-absence matrices taken from the ecology literature. These matrices (together with the original references) can be readily obtained at the link [http://wikieducator.org/Null_Model_Data]. Spatial pattern datasets can be also found at the following online repositories: the Geographic Information System repository [<https://freegisdata.rtwilson.com/>], the National Center for Ecological Analysis and Synthesis repository [<https://www.nceas.ucsb.edu/scicomp/data>].

A.2.3. Ecological interaction networks

Mutualistic network datasets can be found at the following online repositories: Web of Life [<http://www.web-of-life.es/>], Interaction Web DataBase [<https://www.nceas.ucsb.edu/interactionweb/resources.html>]. **Host-parasite** datasets can be found at the following online repositories: Interaction Web DataBase [<https://www.nceas.ucsb.edu/interactionweb/resources.html>]. **Food-web** network datasets can be found at the following online repositories: Web of Life [<http://www.web-of-life.es/>], Interaction Web DataBase [<https://www.nceas.ucsb.edu/interactionweb/resources.html>].

Appendix B. Network-based metrics of node importance in bipartite networks

In the following, we review some of the metrics that have been proposed in the literature to quantify node importance – in particular, those that are relevant to Fig. 26. We refer the interested reader to [291,397] for extensive reviews on centrality metrics and network-based ranking algorithms. Based on the methods and results in [71,72,335], in this review, centrality metrics for bipartite networks have been compared with respect to their ability to rank nodes by their structural importance (Section 6.1.2). We emphasize that centrality metrics are routinely used for different purposes (like the identification of influential nodes for diffusion processes [397], the identification of expert-selected significant nodes [291], the prediction of the nodes' future popularity [280]), and the results presented in Section 6.1.2 only refer to one particular application.

Degree. The degree k_i and k_α of row-node i and column-node α can be considered as the simplest possible centrality metrics in bipartite networks: the two classes of nodes are ranked in order of decreasing degree, under the assumption that a node is central if it has many connections [11].

Closeness. The closeness centrality score c_i of a row-node i is defined as [398]

$$c_i = \frac{M + 2(N - 1)}{d_i}, \quad (\text{B.1})$$

where $d_i = \sum_j D_{ij} + \sum_\alpha D_{i\alpha}$ is the average distance of i from all the other nodes, and D_{ij} denotes the shortest-path distance [11] between i and j . Analogously, for a column-node α , the closeness centrality c_α is defined as [398]

$$c_\alpha = \frac{M + 2(M - 1)}{d_\alpha}, \quad (\text{B.2})$$

where $d_\alpha = \sum_i D_{i\alpha} + \sum_\beta D_{\beta\alpha}$. We refer to [398] for a detailed motivation of the normalization factor, and to the package NetworkX for the implementation [https://networkx.github.io/documentation/networkx-1.10/reference/generated/networkx.algorithms.bipartite.closeness_centrality.html].

Betweenness centrality. The betweenness centrality score of a row-node i (column-node α) is defined as the sum over all pairs of nodes of the share of shortest paths that pass through node i (node α). We refer to [398] for a discussion about how to properly normalize the betweenness score in bipartite networks, and to the package NetworkX for the implementation [https://networkx.github.io/documentation/stable/reference/algorithms/generated/networkx.algorithms.bipartite.centrality.betweenness_centrality.html].

Eigenvector centrality. The vector of eigenvector centrality scores is given by the leading eigenvector of the network's adjacency matrix A [399]. This definition implies that in a bipartite network, the score of a row-node (column-node) is proportional to the scores of its neighbors' column-nodes (row-nodes) [398].

Google's pagerank. Google's PageRank [291,335,400] is defined as the leading eigenvector of the $S \times S$ matrix

$$G = c P + \frac{1 - c}{S} \quad (\text{B.3})$$

where $P_{ij} = A_{ij}/k_i$. The parameter c is referred to as teleportation parameter or damping factor [291]; the results shown in Fig. 26 were obtained by Domínguez-García and Muñoz [71] with $c = 0.999$.

Contribution to nestedness. The individual nodes' contribution to nestedness can be computed in various ways [28,146]. The nestedness contributions used by Saavedra et al. [28] were described in Section 4.3.1. Alternatively, one can take the perspective of overlap-based nestedness metrics, and use them to estimate the individual nodes' contribution. For instance, based on the JDM-NODF metric (Eq. (19)), one can assess the contribution of node i to the overall nestedness as specified by Eq. (20); this is the individual contribution to nestedness that was used by Domínguez-García and Muñoz [71] to obtain the results reported in Fig. 26.

Fitness-complexity algorithm and its reversed variant. The algorithm and some of its variants are detailedly described in Section 3.3.2. Here, we only stress that the original algorithm ranks the row-nodes and column-nodes in order of decreasing fitness and complexity score, respectively. One can also exchange the roles of row-nodes and column-nodes. In a country–product network, this corresponds to assign a fitness score to products and a complexity score to countries; countries and products are, therefore, ranked in order of increasing complexity and increasing fitness, respectively. We refer to the corresponding algorithm as reversed fitness-complexity.

Both algorithms (fitness-complexity algorithm and reversed fitness-complexity) are considered in Fig. 26. They are labeled as "MUS" and "MUSrev" in Fig. 26, as the fitness-complexity algorithm was relabeled as "MUtualistic Species RANKing" algorithm (MusRank) in [71]. As the fitness-complexity and MusRank algorithm are mathematically equivalent, to prevent ambiguity, we always referred to the algorithm as the fitness-complexity algorithm throughout this review.

References

- [1] U. Brandes, G. Robins, A. McCranie, S. Wasserman, What is network science? *Netw. Sci.* 1 (1) (2013) 1–15.
- [2] F.R. Chung, F.C. Graham, *Spectral Graph Theory*, no. 92, American Mathematical Soc., 1997.
- [3] S. Wasserman, K. Faust, *Social Network Analysis: Methods and Applications*, vol. 8, Cambridge University Press, 1994.
- [4] M.O. Jackson, *Social and Economic Networks*, Princeton University Press, 2010.
- [5] D.J.D.S. Price, Networks of scientific papers, *Science* (1965) 510–515.
- [6] H. Small, Co-citation in the scientific literature: A new measure of the relationship between two documents, *J. Am. Soc. Inf. Sci.* 24 (4) (1973) 265–269.
- [7] G. Chartrand, O.R. Oellermann, *Applied and Algorithmic Graph Theory*, vol. 993, McGraw-Hill New York, 1993.
- [8] P. Jordano, Patterns of mutualistic interactions in pollination and seed dispersal: connectance, dependence asymmetries, and coevolution, *Amer. Nat.* 129 (5) (1987) 657–677.
- [9] J. Bascompte, P. Jordano, *Mutualistic Networks*, Princeton University Press, 2013.
- [10] A.-L. Barabási, M. Pósfai, *Network Science*, Cambridge University Press, 2016.
- [11] M. Newman, *Networks: An Introduction*, Oxford university press, 2010.
- [12] G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology*, Oxford University Press, 2007.
- [13] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, *Science* 286 (5439) (1999) 509–512.
- [14] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, *Phys. Rev. Lett.* 86 (14) (2001) 3200.
- [15] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, Epidemic processes in complex networks, *Rev. Modern Phys.* 87 (3) (2015) 925.
- [16] R. Albert, H. Jeong, A.-L. Barabási, Error and attack tolerance of complex networks, *Nature* 406 (6794) (2000) 378.
- [17] Y. Moreno, A.F. Pacheco, Synchronization of kuramoto oscillators in scale-free networks, *Europhys. Lett.* 68 (4) (2004) 603.
- [18] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, C. Zhou, Synchronization in complex networks, *Phys. Rep.* 469 (3) (2008) 93–153.
- [19] E. Hultén, *Outline of the History of Arctic and Boreal Biota During the Quaternary Period*, Bokförlag Aktiebolaget Thule, 1937.
- [20] P.J. Darlington, *Zoogeography*, John Wiley, New York, 1957.
- [21] B.D. Patterson, W. Atmar, Nested subsets and the structure of insular mammalian faunas and archipelagos, *Biol. J. Linnean Soc.* 28 (1–2) (1986) 65–82.
- [22] W. Atmar, B.D. Patterson, The measure of order and disorder in the distribution of species in fragmented habitat, *Oecologia* 96 (3) (1993) 373–382.
- [23] J. Bascompte, P. Jordano, C.J. Melián, J.M. Olesen, The nested assembly of plant-animal mutualistic networks, *Proc. Natl. Acad. Sci.* 100 (16) (2003) 9383–9387.
- [24] A. Tacchella, M. Cristelli, G. Caldarelli, A. Gabrielli, L. Pietronero, A new metrics for countries' fitness and products' complexity, *Sci. Rep.* 2 (2012) 723.
- [25] M.D. König, C.J. Tessone, Y. Zenou, Nestedness in networks: A theoretical model and some applications, *Theor. Econ.* 9 (3) (2014) 695–752.
- [26] F. Saracco, R. Di Clemente, A. Gabrielli, T. Squartini, Detecting early signs of the 2007–2008 crisis in the world trade, *Sci. Rep.* 6 (2016).
- [27] S. Saavedra, F. Reed-Tsochas, B. Uzzi, A simple model of bipartite cooperation for ecological and organizational networks, *Nature* 457 (7228) (2009) 463.
- [28] S. Saavedra, D.B. Stouffer, B. Uzzi, J. Bascompte, Strong contributors to network persistence are the most vulnerable to extinction, *Nature* 478 (7368) (2011) 233.
- [29] S. Bustos, C. Gomez, R. Hausmann, C.A. Hidalgo, The dynamics of nestedness predicts the evolution of industrial ecosystems, *PLoS One* 7 (11) (2012) e49393.
- [30] A. Garas, C. Rozenblat, F. Schweitzer, Economic specialization and the nested bipartite network of city-firm relations, in: *Multiplex and Multilevel Networks*, 2018, pp. 74–83.
- [31] K. Soramäki, M.L. Bech, J. Arnold, R.J. Glass, W.E. Beyeler, The topology of interbank payment flows, *Physica A* 379 (1) (2007) 317–333.
- [32] J. Borge-Holthoefer, R.A. Baños, C. Gracia-Lázaro, Y. Moreno, Emergence of consensus as a modular-to-nested transition in communication dynamics, *Sci. Rep.* 7 (2017).
- [33] S. Suweis, F. Simini, J.R. Banavar, A. Maritan, Emergence of structural and dynamical properties of ecological mutualistic networks, *Nature* 500 (7463) (2013) 449–452.
- [34] S. Valverde, J. Piñero, B. Corominas-Murtra, J. Montoya, L. Joppa, R. Solé, The architecture of mutualistic networks as an evolutionary spandrel, *Nat. Ecol. Evol.* 2 (1) (2018) 94.
- [35] D.S. Maynard, C.A. Serván, S. Allesina, Network spandrels reflect ecological assembly, *Ecol. Lett.* 21 (3) (2018) 324–334.
- [36] M.D. König, C.J. Tessone, Network evolution based on centrality, *Phys. Rev. E* 84 (5) (2011) 056108.
- [37] S. Allesina, S. Tang, Stability criteria for complex ecosystems, *Nature* 483 (2012) 205–208.
- [38] R.P. Rohr, S. Saavedra, J. Bascompte, On the structural stability of mutualistic systems, *Science* 345 (6195) (2014) 1253497.
- [39] M.A. Fortuna, D.B. Stouffer, J.M. Olesen, P. Jordano, D. Mouillot, B.R. Krasnov, R. Poulin, J. Bascompte, Nestedness versus modularity in ecological networks: two sides of the same coin? *J. Anim. Ecol.* 79 (4) (2010) 811–817.
- [40] A. Solé-Ribalta, C.J. Tessone, M.S. Mariani, J. Borge-Holthoefer, Revealing in-block nestedness: Detection and benchmarking, *Phys. Rev. E* 97 (6) (2018) 062302.
- [41] S.H. Lee, et al., Network nestedness as generalized core-periphery structures, *Phys. Rev. E* 93 (2) (2016) 022306.
- [42] G. Strona, F. Stefani, P. Galli, S. Fattorini, A protocol to compare nestedness among submatrices, *Popul. Ecol.* 55 (1) (2013) 227–239.
- [43] A. Grimm, C.J. Tessone, Analysing the sensitivity of nestedness detection methods, *Appl. Netw. Sci.* 2 (1) (2017) 37.
- [44] S. Kojaku, N. Masuda, Core-periphery structure requires something else in the network, *New J. Phys.* 20 (4) (2018) 043012.
- [45] A. Tacchella, D. Mazzilli, L. Pietronero, A dynamical systems approach to gross domestic product forecasting, *Nat. Phys.* (2018) 1.
- [46] J. Bascompte, P. Jordano, Plant-animal mutualistic networks: the architecture of biodiversity, *Annu. Rev. Ecol. Evol. Syst.* 38 (2007) 567–593.
- [47] T.C. Ings, J.M. Montoya, J. Bascompte, N. Blüthgen, L. Brown, C.F. Dormann, F. Edwards, D. Figueroa, U. Jacob, J.I. Jones, et al., Ecological networks—beyond food webs, *J. Anim. Ecol.* 78 (1) (2009) 253–269.
- [48] W. Ulrich, M. Almeida-Neto, N.J. Gotelli, A consumer's guide to nestedness analysis, *Oikos* 118 (1) (2009) 3–17.
- [49] C.F. Dormann, J. Fründ, N. Blüthgen, B. Gruber, Indices, graphs and null models: analyzing bipartite ecological networks, *Open Ecol. J.* 2 (1) (2009).
- [50] U. Bastolla, M.A. Fortuna, A. Pascual-García, A. Ferrera, B. Luque, J. Bascompte, The architecture of mutualistic networks minimizes competition and increases biodiversity, *Nature* 458 (7241) (2009) 1018.
- [51] F. Saracco, R. Di Clemente, A. Gabrielli, L. Pietronero, From innovation to diversification: a simple competitive model, *PLoS One* 10 (11) (2015) e0140420.
- [52] R. Hausmann, C.A. Hidalgo, The network structure of economic output, *J. Econ. Growth* 16 (4) (2011) 309–342.

- [53] N. Masuda, H. Miwa, N. Konno, Analysis of scale-free networks based on a threshold graph with intrinsic vertex weights, *Phys. Rev. E* 70 (3) (2004) 036124.
- [54] A. Hagberg, P.J. Swart, D.A. Schult, Designing threshold networks with given structural and dynamical properties, *Phys. Rev. E* 74 (5) (2006) 056116.
- [55] P. Holme, J. Saramäki, Temporal networks, *Phys. Rep.* 519 (3) (2012) 97–125.
- [56] P. Holme, Modern temporal network theory: a colloquium, *Eur. Phys. J. B* 88 (9) (2015) 234.
- [57] S. Boccaletti, G. Bianconi, R. Criado, C.I. Del Genio, J. Gómez-Gardenes, M. Romance, I. Sendina-Nadal, Z. Wang, M. Zanin, The structure and dynamics of multilayer networks, *Phys. Rep.* 544 (1) (2014) 1–122.
- [58] M. Kivelä, A. Arenas, M. Barthelemy, J.P. Gleeson, Y. Moreno, M.A. Porter, Multilayer networks, *J. Complex Netw.* 2 (3) (2014) 203–271.
- [59] M.J. Straka, G. Caldarelli, T. Squartini, F. Saracco, From ecology to finance (and back?): Recent advancements in the analysis of bipartite networks, *J. Stat. Phys.* (2018) <https://doi.org/10.1007/s10955-018-2039-4>.
- [60] N.J. Gotelli, Null model analysis of species co-occurrence patterns, *Ecology* 81 (9) (2000) 2606–2621.
- [61] W. Ulrich, N.J. Gotelli, Null model analysis of species nestedness patterns, *Ecology* 88 (7) (2007) 1824–1831.
- [62] N.J. Gotelli, W. Ulrich, Statistical challenges in null model analysis, *Oikos* 121 (2) (2012) 171–180.
- [63] G. Strona, S. Fattorini, On the methods to assess significance in nestedness analyses, *Theory Biosci.* 133 (3–4) (2014) 179–186.
- [64] G. Strona, W. Ulrich, N.J. Gotelli, Bi-dimensional null model analysis of presence-absence binary matrices, *Ecology* 99 (1) (2018) 103–115.
- [65] E.F. Connor, D. Simberloff, The assembly of species communities: chance or competition? *Ecology* 60 (6) (1979) 1132–1140.
- [66] B.G. Jonsson, A null model for randomization tests of nestedness in species assemblages, *Oecologia* 127 (3) (2001) 309–313.
- [67] L.N. Joppa, J.M. Montoya, S. Vicente, J. Sanderson, S.L. Pimm, et al., On nestedness in ecological networks, *Evol. Ecol. Res.* 12 (2010) 35–46.
- [68] A. Krishna, P.R. Guimaraes Jr., P. Jordano, J. Bascompte, A neutral-niche theory of nestedness in mutualistic networks, *Oikos* 117 (11) (2008) 1609–1618.
- [69] R. Kadmon, Nested species subsets and geographic isolation: a case study, *Ecology* 76 (2) (1995) 458–465.
- [70] M.A. Rodríguez-Gironés, L. Santamaría, A new algorithm to calculate the nestedness temperature of presence-absence matrices, *J. Biogeogr.* 33 (5) (2006) 924–935.
- [71] V. Domínguez-García, M.A. Muñoz, Ranking species in mutualistic networks, *Sci. Rep.* 5 (2015) 8182.
- [72] M.S. Mariani, A. Vidmer, M. Medo, Y.-C. Zhang, Measuring economic complexity of countries and products: which metric to use? *Eur. Phys. J. B* 88 (11) (2015) 293.
- [73] J.-H. Lin, C. Tessone, M. Mariani, Nestedness maximization in complex networks through the fitness-complexity algorithm, *Entropy* 20 (10) (2018) 768.
- [74] R.R. Cook, J.F. Quinn, The influence of colonization in nested species subsets, *Oecologia* 102 (4) (1995) 413–424.
- [75] P.F. Sale, Overlap in resource use, and interspecific competition, *Oecologia* 17 (3) (1974) 245–256.
- [76] R.F. Inger, R.K. Colwell, Organization of contiguous communities of amphibians and reptiles in Thailand, *Ecol. Monograph* 47 (3) (1977) 229–253.
- [77] M.E. Gilpin, J.M. Diamond, Factors contributing to non-randomness in species co-occurrences on islands, *Oecologia* 52 (1) (1982) 75–84.
- [78] F. Hering, Nested bipartite graphs, *Israel J. Math.* 9 (3) (1971) 403–417.
- [79] N.V. Mahadev, U.N. Peled, *Threshold Graphs and Related Topics*, vol. 56, Elsevier, 1995.
- [80] D. Medan, R.P. Perazzo, M. Devoto, E. Burgos, M.G. Zimmermann, H. Ceva, A.M. Delbue, Analysis and assembling of network structure in mutualistic systems, *J. Theoret. Biol.* 246 (3) (2007) 510–521.
- [81] M. Bardoscia, G. De Luca, G. Livan, M. Marsili, C.J. Tessone, The social climbing game, *J. Stat. Phys.* 151 (3–4) (2013) 440–457.
- [82] F.K. Bell, D. Cvetković, P. Rowlinson, S.K. Simić, Graphs for which the least eigenvalue is minimal, II, *Linear Algebra Appl.* 429 (8–9) (2008) 2168–2179.
- [83] P.P. Staniczenko, J.C. Kopp, S. Allesina, The ghost of nestedness in ecological networks, *Nature Commun.* 4 (2013) 1391.
- [84] W.-M. Song, T. Di Matteo, T. Aste, Nested hierarchies in planar graphs, *Discrete Appl. Math.* 159 (17) (2011) 2135–2146.
- [85] W.-M. Song, T. Di Matteo, T. Aste, Hierarchical information clustering by means of topologically embedded graphs, *PLoS One* 7 (3) (2012) e31929.
- [86] S. Földes, P.L. Hammer, On split graphs and some related questions, *Probl. Combin. Théorie Graphes* (1976) 138–140.
- [87] Z.A. Chernyak, About recognizing (a, b)-classes of polar graphs, *Discrete Math.* 62 (2) (1986) 133–138.
- [88] R.D. Luce, A.D. Perry, A method of matrix analysis of group structure, *Psychometrika* 14 (2) (1949) 95–116.
- [89] C. Berge, Two theorems in graph theory, *Proc. Natl. Acad. Sci.* 43 (9) (1957) 842–844.
- [90] D. Cvetković, P. Rowlinson, The largest eigenvalue of a graph: A survey, *Linear and multilinear algebra* 28 (1–2) (1990) 3–33.
- [91] R.A. Brualdi, A.J. Hoffman, On the spectral radius of (0, 1)-matrices, *Linear Algebra Appl.* 65 (1985) 133–146.
- [92] V. Chvátal, P.L. Hammer, Aggregation of inequalities in integer programming, *Ann. Discrete Math.* 1 (1977) 145–162.
- [93] N.T. Wheelwright, G.H. Orians, Seed dispersal by animals: contrasts with pollen dispersal, problems of terminology, and constraints on coevolution, *Amer. Nat.* 119 (3) (1982) 402–413.
- [94] R.M. May, Will a large complex system be stable? *Nature* 238 (5364) (1972) 413.
- [95] M. Kondoh, S. Kato, Y. Sakato, Food webs are built up with nested subwebs, *Ecology* 91 (11) (2010) 3123–3130.
- [96] T.M. Lewinsohn, P. Inácio Prado, P. Jordano, J. Bascompte, J. M. Olesen, Structure in plant-animal interaction assemblages, *Oikos* 113 (1) (2006) 174–184.
- [97] S.P. Graham, H.K. Hassan, N.D. Burkett-Cadena, C. Guyer, T.R. Unnasch, Nestedness of ectoparasite-vertebrate host networks, *PLoS One* 4 (11) (2009) e7873.
- [98] S. Pilosof, M.A. Fortuna, J.-F. Cosson, M. Galan, C. Kittipong, A. Ribas, E. Segal, B.R. Krasnov, S. Morand, J. Bascompte, Host-parasite network structure is associated with community-level immunogenetic diversity, *Nature Commun.* 5 (2014) 5172.
- [99] C. McQuaid, N. Britton, Host-parasite nestedness: a result of co-evolving trait-values, *Ecol. Complex.* 13 (2013) 53–59.
- [100] G. Gaulier, S. Zignago, BacI: International Trade Database at the Product-Level (the 1994–2007 Version), CEPII Working Paper, 2010–23, 2010.
- [101] R.M. May, S.A. Levin, G. Sugihara, Complex systems: Ecology for bankers, *Nature* 451 (7181) (2008) 893–895.
- [102] T. Squartini, G. Fagiolo, D. Garlaschelli, Randomizing world trade. I. A binary network analysis, *Phys. Rev. E* 84 (4) (2011) 046117.
- [103] T. Squartini, G. Fagiolo, D. Garlaschelli, Randomizing world trade. II. A weighted network analysis, *Phys. Rev. E* 84 (4) (2011) 046118.
- [104] Z. Zhu, F. Cerina, A. Chessa, G. Caldarelli, M. Riccaboni, The rise of China in the international trade network: a community core detection approach, *PLoS One* 9 (8) (2014) e105496.
- [105] L. De Benedictis, L. Tajoli, The world trade network, *World Econ.* 34 (8) (2011) 1417–1454.
- [106] A. Akerman, A.L. Seim, The global arms trade network 1950–2007, *J. Comp. Econ.* 42 (3) (2014) 535–551.
- [107] Z.-M. Ren, A. Zeng, Y.-C. Zhang, Bridging nestedness and economic complexity in multilayer international trading networks, 2019, in preparation.
- [108] L.G. Alves, G. Mangioni, I. Cingolani, F.A. Rodrigues, P. Panzarasa, Y. Moreno, The nested structural organization of the worldwide trade multi-layer network, *Scientific Reports* 9 (1) (2019) 2866.

- [109] C. Piccardi, L. Tajoli, Complexity, centralization, and fragility in economic networks, *PLoS one* 13 (11) (2018) e0208265.
- [110] C.A. Hidalgo, R. Hausmann, The building blocks of economic complexity, *Proc. Natl. Acad. Sci.* 106 (26) (2009) 10570–10575.
- [111] L. Ermann, D.L. Shepelyansky, Ecological analysis of world trade, *Phys. Lett. A* 377 (3) (2013) 250–256.
- [112] F. Saracco, R. Di Clemente, A. Gabrielli, T. Squartini, Randomizing bipartite networks: the case of the world trade web, *Sci. Rep.* 5 (2015) 10595.
- [113] O. Angelini, T. Di Matteo, Complexity of products: the effect of data regularisation, *Entropy* 20 (11) (2018) 814.
- [114] E. Pugliese, A. Zaccaria, L. Pietronero, On the convergence of the fitness-complexity algorithm, *Eur. Phys. J. Spec. Top.* 225 (10) (2016) 1893–1911.
- [115] R.-J. Wu, G.-Y. Shi, Y.-C. Zhang, M.S. Mariani, The mathematics of non-linear metrics for nested networks, *Physica A* 460 (2016) 254–269.
- [116] V. Stojkoski, Z. Utkovski, L. Kocarev, The impact of services on economic complexity: Service sophistication as route for economic growth, *PLoS One* 11 (8) (2016) e0161633.
- [117] V. Servedio, P. Buttà, D. Mazzilli, A. Tacchella, L. Pietronero, A new and stable estimation method of country economic fitness and product complexity, *Entropy* 20 (10) (2018) 783.
- [118] D. Ricardo, *On the Principles of Political Economy and Taxation*, John Murray, London, 1817.
- [119] J. Imbs, R. Wacziarg, Stages of diversification, *Amer. Econ. Rev.* 93 (1) (2003) 63–86.
- [120] B. Balassa, Trade liberalisation and “revealed” comparative advantage 1, *Manch. Sch.* 33 (2) (1965) 99–123.
- [121] M. Cristelli, A. Tacchella, L. Pietronero, The heterogeneous dynamics of economic complexity, *PLoS One* 10 (2) (2015) e0117174.
- [122] M.C.A. Cristelli, A. Tacchella, M.Z. Cader, K.I. Roster, L. Pietronero, On the predictability of growth, World Bank Policy Research Working Paper No. 8117, 2017, Available at SSRN: <https://ssrn.com/abstract=3006151>.
- [123] I. McCarthy, A. Anagnostou, The impact of outsourcing on the transaction costs and boundaries of manufacturing, *Int. J. Prod. Econ.* 88 (1) (2004) 61–71.
- [124] B. Uzzi, The sources and consequences of embeddedness for the economic performance of organizations: The network effect, *Am. Sociol. Rev.* (1996) 674–698.
- [125] L. Hernández, A. Vignes, S. Saba, Trust or robustness? An ecological approach to the study of auction and bilateral markets, *PLoS One* 13 (5) (2018) e0196206.
- [126] F. Kyriakopoulos, S. Thurner, C. Puhr, S.W. Schmitz, Network and eigenvalue analysis of financial transaction networks, *Eur. Phys. J. B* 71 (4) (2009) 523.
- [127] B. Craig, G. Von Peter, Interbank tiering and money center banks, *J. Financ. Intermed.* 23 (3) (2014) 322–347.
- [128] D. Fricke, T. Lux, Core-periphery structure in the overnight money market: evidence from the e-mid trading platform, *Comput. Econ.* 45 (3) (2015) 359–395.
- [129] P. Barucca, F. Lillo, Disentangling bipartite and core-periphery structure in financial networks, *Chaos Solitons Fractals* 88 (2016) 244–253.
- [130] S. Kojaku, G. Cimini, G. Caldarelli, N. Masuda, Structural changes in the interbank market across the financial crisis from multiple core-periphery analysis, *J. Netw. Theory in Financ.* 4 (3) (2018) 33–51.
- [131] J. Gao, T. Zhou, Quantifying China’s regional economic complexity, *Physica A* 492 (2018) 1591–1603.
- [132] H.J. Leavitt, Some effects of certain communication patterns on group performance, *J. Abnorm. Soc. Psychol.* 46 (1) (1951) 38.
- [133] S.P. Borgatti, A. Mehra, D.J. Brass, G. Labianca, Network analysis in the social sciences, *Science* 323 (5916) (2009) 892–895.
- [134] B. Kogut, U. Zander, Knowledge of the firm, combinative capabilities, and the replication of technology, *Organ. Sci.* 3 (3) (1992) 383–397.
- [135] B. Kogut, U. Zander, What firms do? Coordination, identity, and learning, *Organ. Sci.* 7 (5) (1996) 502–518.
- [136] D. Krackhardt, R.N. Stern, Informal networks and organizational crises: An experimental simulation, *Soc. Psychol. Q.* (1988) 123–140.
- [137] D. Krackhardt, Informal networks: the company behind the chart, *Harv. Bus. Rev.* 71 (1993) 104–111.
- [138] O.N. Rank, Formal structures and informal networks: Structural analysis in organizations, *Scand. J. Manag.* 24 (2) (2008) 145–161.
- [139] S. Valverde, R.V. Solé, Self-organization versus hierarchy in open-source social networks, *Phys. Rev. E* 76 (4) (2007) 046118.
- [140] M.A. Fortuna, J.A. Bonachela, S.A. Levin, Evolution of a modular software network, *Proc. Natl. Acad. Sci.* 108 (50) (2011) 19985–19989.
- [141] M. Bastos, C. Piccardi, M. Levy, N. McRoberts, M. Lubell, Core-periphery or decentralized? Topological shifts of specialized information on Twitter, *Social Networks* 52 (2018) 282–293.
- [142] E.M. Rogers, *Diffusion of Innovations*, Simon and Schuster, 2010.
- [143] J.M. Kamilar, Q.D. Atkinson, Patterns of cultural nestedness in humans, chimpanzees, and orangutans, *Am. J. Phys. Anthropol.* 153 (2014) 154–155.
- [144] M. Cantor, M.M. Pires, F.M. Marquitti, R.L. Raimundo, E. Sebastián-González, P.P. Coltri, S.I. Perez, D.R. Barneche, D.Y. Brandt, K. Nunes, et al., Nestedness across biological scales, *PLoS One* 12 (2) (2017) e0171691.
- [145] D.S. Bassett, N.F. Wymbs, M.P. Rombach, M.A. Porter, P.J. Mucha, S.T. Grafton, Task-based core-periphery organization of human brain dynamics, *PLoS Comput. Biol.* 9 (9) (2013) e1003171.
- [146] S. Johnson, V. Domínguez-García, M.A. Muñoz, Factors determining nestedness in complex networks, *PLoS One* 8 (9) (2013) e74025.
- [147] W.W. Zachary, An information flow model for conflict and fission in small groups, *J. Anthropol. Res.* 33 (4) (1977) 452–473.
- [148] P.M. Gleiser, L. Danon, Community structure in jazz, *Adv. Complex Syst.* 6 (04) (2003) 565–573.
- [149] S. Jonhson, V. Domínguez-García, M.A. Muñoz, Factors determining nestedness in complex networks, *PLoS One* 8 (9) (2013) e74025.
- [150] M. Almeida-Neto, P. Guimarães, P.R. Guimarães, R.D. Loyola, W. Ulrich, A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement, *Oikos* 117 (8) (2008) 1227–1239.
- [151] M. Almeida-Neto, W. Ulrich, A straightforward computational approach for measuring nestedness using quantitative matrices, *Environ. Model. Softw.* 26 (2) (2011) 173–178.
- [152] A. Cutler, Nested faunas and extinction in fragmented habitats, *Conserv. Biol.* 5 (4) (1991) 496–504.
- [153] M.V. Lomolino, Investigating causality of nestedness of insular communities: selective immigrations or extinctions? *J. Biogeography* 23 (5) (1996) 699–703.
- [154] R.A. Brundt, J.G. Sanderson, Nested species subsets, gaps, and discrepancy, *Oecologia* 119 (2) (1999) 256–264.
- [155] L. Boltzmann, Weitere studien fiber das wsxmgleichgewicht unter gasmolekilen, *Wien. Ber.* 66 (275) (1872) 3.
- [156] C.E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.* 27 (1948) 623–656.
- [157] P.R. Guimarães, P. Guimaraes, Improving the analyses of nestedness for large sets of matrices, *Environ. Model. Softw.* 21 (10) (2006) 1512–1513.
- [158] G. Corso, N.F. Britton, Nestedness and τ -temperature in ecological networks, *Ecol. Complex.* 11 (2012) 137–143.
- [159] D.H. Wright, J.H. Reeves, On the meaning and measurement of nestedness of species assemblages, *Oecologia* 92 (3) (1992) 416–428.
- [160] S.J. Beckett, C.A. Boulton, H.T. Williams, FALCON: a software package for analysis of nestedness in bipartite networks, *F1000Research* 3 (2014).
- [161] A. Bhattacharya, S. Friedland, U.N. Peled, On the first eigenvalue of bipartite graphs, *Electron. J. Combin.* 15 (1) (2008) 144.
- [162] W.J. Stewart, *Introduction to the Numerical Solution of Markov Chains*, Princeton University Press, 1994.
- [163] M. Molloy, B. Reed, A critical point for random graphs with a given degree sequence, *Random Struct. Algorithms* 6 (2–3) (1995) 161–180.
- [164] M.E. Newman, S.H. Strogatz, D.J. Watts, Random graphs with arbitrary degree distributions and their applications, *Phys. Rev. E* 64 (2) (2001) 026118.

- [165] S. Maslov, K. Sneppen, Specificity and stability in topology of protein networks, *Science* 296 (5569) (2002) 910–913.
- [166] F. Chung, L. Lu, Connected components in random graphs with given expected degree sequences, *Ann. Comb.* 6 (2) (2002) 125–145.
- [167] T. Squartini, D. Garlaschelli, Analytical maximum-likelihood method to detect patterns in real networks, *New J. Phys.* 13 (8) (2011) 083001.
- [168] G. Parisi, Statistical field theory, 1988.
- [169] C.P. Borras, L. Hernández, Y. Moreno, Breaking the spell of nestedness, 2017, arXiv preprint [arXiv:1711.03134](https://arxiv.org/abs/1711.03134).
- [170] K.O. Winemiller, E.R. Pianka, Organization in natural assemblages of desert lizards and tropical fishes, *Ecol. Monograph* 60 (1) (1990) 27–55.
- [171] J.M. Diamond, M.E. Gilpin, Examination of the “null” model of Connor and Simberloff for species co-occurrences on islands, *Oecologia* 52 (1) (1982) 64–74.
- [172] J.G. Sanderson, M.P. Moulton, R.G. Selfridge, Null matrices and the analysis of species co-occurrences, *Oecologia* 116 (1–2) (1998) 275–283.
- [173] N.J. Gotelli, G.L. Entsminger, Swap and fill algorithms in null model analysis: rethinking the knight’s tour, *Oecologia* 129 (2) (2001) 281–291.
- [174] I. Miklós, J. Podani, Randomization of presence-absence matrices: comments and new algorithms, *Ecology* 85 (1) (2004) 86–92.
- [175] G. Strona, D. Nappo, F. Boccacci, S. Fattorini, J. San-Miguel-Ayanz, A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals, *Nature Commun.* 5 (2014) 4114.
- [176] C.J. Carstens, Proof of uniform sampling of binary matrices with fixed row sums and column sums for the fast curveball algorithm, *Phys. Rev. E* 91 (4) (2015) 042812.
- [177] B.K. Fosdick, D.B. Larremore, J. Nishimura, J. Ugander, Configuring random graph models with fixed degree sequences, *SIAM Rev.* 60 (2) (2018) 315–355.
- [178] B.D. Coleman, M.A. Mares, M.R. Willig, Y.-H. Hsieh, Randomness, area, and species richness, *Ecology* 63 (4) (1982) 1121–1133.
- [179] G.R. Graves, N.J. Gotelli, Assembly of avian mixed-species flocks in Amazonia, *Proc. Natl. Acad. Sci.* 90 (4) (1993) 1388–1391.
- [180] L.G. Abele, W.K. Patton, The size of coral heads and the community biology of associated decapod crustaceans, *J. Biogeogr.* (1976) 35–47.
- [181] G. Casiraghi, V. Nanumyan, I. Scholtes, F. Schweitzer, Generalized hypergeometric ensembles: statistical hypothesis testing in complex networks, 2016, arXiv preprint [arXiv:1607.02441](https://arxiv.org/abs/1607.02441).
- [182] L. Stone, A. Roberts, The checkerboard score and species distributions, *Oecologia* 85 (1) (1990) 74–79.
- [183] W. Ulrich, N. J. Gotelli, Disentangling community patterns of nestedness and species co-occurrence, *Oikos* 116 (12) (2007) 2053–2061.
- [184] J.E. Moore, R.K. Swihart, Toward ecologically explicit null models of nestedness, *Oecologia* 152 (4) (2007) 763–777.
- [185] D.P. Vázquez, M.A. Aizen, Community-wide patterns of specialization in plant-pollinator interactions revealed by null models, in: *Plant-Pollinator Interactions: From Specialization to Generalization*, University of Chicago Press Chicago, 2006, pp. 200–219.
- [186] D.P. Vázquez, On the parsimony of null models of plant-pollinator networks, *PLoS Biol.* (2007) <http://biology.plosjournals.org/perlserv?request=read-response&doi=10.1371/journal.pbio.0050031#r1695>.
- [187] M. Stang, P.G.L. Klinkhamer, N.M. Waser, E. van der Meijden, Both constraints and abundance are needed to explain the topology of mutualistic networks, *PLoS Biol.* (2007) <http://journals.plos.org/plosbiology/article/comment?id=10.1371/annotation/79298be1-fa22-40fb-abc6-c11d5569e246>.
- [188] R. Mastrandrea, T. Squartini, G. Fagiolo, D. Garlaschelli, Enhanced reconstruction of weighted networks from strengths and degrees, *New J. Phys.* 16 (4) (2014) 043022.
- [189] T. Squartini, G. Cimini, A. Gabrielli, D. Garlaschelli, Network reconstruction via density sampling, *Appl. Netw. Sci.* 2 (1) (2017) 3.
- [190] T. Squartini, I. Van Lelyveld, D. Garlaschelli, Early-warning signals of topological collapse in interbank networks, *Sci. Rep.* 3 (2013) 3357.
- [191] J. Fischer, D.B. Lindenmayer, Treating the nestedness temperature calculator as a “black box” can lead to false conclusions, *Oikos* 99 (1) (2002) 193–199.
- [192] D. Whitley, A genetic algorithm tutorial, *Stat. Comput.* 4 (2) (1994) 65–85.
- [193] M. Cristelli, A. Gabrielli, A. Tacchella, G. Caldarelli, L. Pietronero, Measuring the intangibles: A metrics for the economic complexity of countries and products, *PLoS One* 8 (8) (2013) e70726.
- [194] P. Berkhin, A survey on pagerank computing, *Internet Math.* 2 (1) (2005) 73–120.
- [195] C. Tu, J. Carr, S. Suweis, A data driven network approach to rank countries production diversity and food specialization, *PLoS One* 11 (11) (2016) e0165941.
- [196] G. Cimini, A. Gabrielli, F.S. Labini, The scientific competitiveness of nations, *PLoS One* 9 (12) (2014) e113470.
- [197] A. Zaccaria, M. Del Vicario, W. Quattrociocchi, A. Scala, L. Pietronero, Poprank: Ranking pages’ impact and userse’ ngagement on facebook, *PLoS One* 14 (1) (2019) e0211038.
- [198] F. Battiston, M. Cristelli, A. Tacchella, L. Pietronero, How metrics for economic complexity are affected by noise, *Complex. Econ.* 3 (1) (2014) 1–22.
- [199] A. Zaccaria, S. Mishra, M. Cader, L. Pietronero, Integrating Services in the Economic Fitness Approach, World Bank Policy Research Working Paper No. 8485, 2018, Available at SSRN: <https://ssrn.com/abstract=3238321>.
- [200] J.M. Olesen, J. Bascompte, Y.L. Dupont, P. Jordano, The modularity of pollination networks, *Proc. Natl. Acad. Sci.* 104 (50) (2007) 19891–19896.
- [201] C. Orsini, M.M. Dankulov, P. Colomer-de Simón, A. Jamakovic, P. Mahadevan, A. Vahdat, K.E. Bassler, Z. Toroczkai, M. Boguñá, G. Caldarelli, S. Fortunato, D. Krioukov, Quantifying randomness in real networks, *Nature Commun.* 6 (2015) 8627.
- [202] D. Simberloff, J.-L. Martin, Nestedness of insular avifaunas: simple summary statistics masking complex species patterns, *Ornis Fenn.* 68 (4) (1991) 178–192.
- [203] S. Johnson, J.J. Torres, J. Marro, M.A. Munoz, Entropic origin of disassortativity in complex networks, *Phys. Rev. Lett.* 104 (10) (2010) 108702.
- [204] S.J. Presley, C.L. Higgins, M.R. Willig, A comprehensive framework for the evaluation of metacommunity structure, *Oikos* 119 (6) (2010) 908–917.
- [205] M.E. Newman, Assortative mixing in networks, *Phys. Rev. Lett.* 89 (20) (2002) 208701.
- [206] G. Abramson, C.A.T. Soto, L. Ona, The role of asymmetric interactions on the effect of habitat destruction in mutualistic networks, *PLoS One* 6 (6) (2011) e21028.
- [207] J. Ugander, B. Karrer, L. Backstrom, C. Marlow, The anatomy of the facebook social graph, 2011, arXiv preprint [arXiv:1111.4503](https://arxiv.org/abs/1111.4503).
- [208] L. Bornmann, F. De Moya Anegón, L. Leydesdorff, Do scientific advancements lean on the shoulders of giants? A bibliometric investigation of the ortega hypothesis, *PLoS One* 5 (10) (2010) e13327.
- [209] M.S. Mariani, M. Medo, F. Lafond, Early identification of important patents: Design and validation of citation network metrics, *Technol. Forecast. Soc. Change* (2018) <http://dx.doi.org/10.1016/j.techfore.2018.01.036>.
- [210] J.G. Foster, D.V. Foster, P. Grassberger, M. Paczuski, Edge direction and the structure of networks, *Proc. Natl. Acad. Sci.* 107 (24) (2010) 10815–10820.
- [211] R. Pastor-Satorras, A. Vázquez, A. Vespignani, Dynamical and correlation properties of the internet, *Phys. Rev. Lett.* 87 (25) (2001) 258701.
- [212] R.J. Nemeth, D.A. Smith, International trade and world-system structure: A multiple network analysis, *Review (Fernand Braudel Center)* 8 (4) (1985) 517–560.
- [213] D.A. Smith, D.R. White, Structure and dynamics of the global economy: network analysis of international trade 1965–1980, *Soc. Forces* 70 (4) (1992) 857–893.
- [214] S.P. Borgatti, M.G. Everett, Models of core/periphery structures, *Social Networks* 21 (4) (2000) 375–395.

- [215] G. Cattani, S. Ferriani, A core/periphery perspective on individual creative performance: Social networks and cinematic achievements in the hollywood film industry, *Organ. Sci.* 19 (6) (2008) 824–844.
- [216] R.F. Betzel, D.S. Bassett, Multi-scale brain networks, *Neuroimage* 160 (2017) 73–83.
- [217] F. Battiston, J. Guillou, M. Chavez, V. Latora, F.D.V. Fallani, Multiplex core-periphery organization of the human connectome, *J. Roy. Soc. Interface* 15 (146) (2018).
- [218] C. Díaz-Castelazo, P.R. Guimaraes, P. Jordano, J.N. Thompson, R.J. Marquis, V. Rico-Gray, Changes of a mutualistic network over time: reanalysis over a 10-year period, *Ecology* 91 (3) (2010) 793–801.
- [219] R.A. Ruggera, P.G. Blendinger, M.D. Gomez, C. Marshak, Linking structure and functionality in mutualistic networks: do core frugivores disperse more seeds than peripheral species? *Oikos* 125 (4) (2016) 541–555.
- [220] M.V. Tomasello, M. Napoletano, A. Garas, F. Schweitzer, The rise and fall of R&D networks, *Ind. Corp. Change* 26 (4) (2017) 617–646.
- [221] P. Csermely, A. London, L.-Y. Wu, B. Uzzi, Structure and dynamics of core/periphery networks, *J. Complex Netw.* 1 (2) (2013) 93–123.
- [222] S. Fortunato, Community detection in graphs, *Phys. Rep.* 486 (3–5) (2010) 75–174.
- [223] S. Fortunato, D. Hric, Community detection in networks: A user guide, *Phys. Rep.* 659 (2016) 1–44.
- [224] P. Rombach, M.A. Porter, J.H. Fowler, P.J. Mucha, Core-periphery structure in networks (revisited), *SIAM Rev.* 59 (3) (2017) 619–646.
- [225] P. Holme, Core-periphery organization of complex networks, *Phys. Rev. E* 72 (4) (2005) 046111.
- [226] J.P. Boyd, W.J. Fitzgerald, M.C. Mahutga, D.A. Smith, Computing continuous core/periphery structures for social relations data with MINRES/SVD, *Social Networks* 32 (2) (2010) 125–137.
- [227] X. Zhang, T. Martin, M.E. Newman, Identification of core-periphery structure in networks, *Phys. Rev. E* 91 (3) (2015) 032803.
- [228] P. Bonacich, Power and centrality: A family of measures, *Am. J. Sociol.* 92 (5) (1987) 1170–1182.
- [229] M. Rosvall, C.T. Bergstrom, Maps of random walks on complex networks reveal community structure, *Proc. Natl. Acad. Sci.* 105 (4) (2008) 1118–1123.
- [230] P.W. Holland, K.B. Laskey, S. Leinhardt, Stochastic blockmodels: First steps, *Social Networks* 5 (2) (1983) 109–137.
- [231] B. Karrer, M.E. Newman, Stochastic blockmodels and community structure in networks, *Phys. Rev. E* 83 (1) (2011) 016107.
- [232] F. Krzakala, C. Moore, E. Mossel, J. Neeman, A. Sly, L. Zdeborová, P. Zhang, Spectral redemption in clustering sparse networks, *Proc. Natl. Acad. Sci.* 110 (52) (2013) 20935–20940.
- [233] M.E. Newman, M. Girvan, Finding and evaluating community structure in networks, *Phys. Rev. E* 69 (2) (2004) 026113.
- [234] F.D. Malliaros, M. Vazirgiannis, Clustering and community detection in directed networks: A survey, *Phys. Rep.* 533 (4) (2013) 95–142.
- [235] M.J. Barber, Modularity and community detection in bipartite networks, *Phys. Rev. E* 76 (6) (2007) 066102.
- [236] A. Lancichinetti, S. Fortunato, F. Radicchi, Benchmark graphs for testing community detection algorithms, *Phys. Rev. E* 78 (4) (2008) 046110.
- [237] A. Lancichinetti, S. Fortunato, Community detection algorithms: a comparative analysis, *Phys. Rev. E* 80 (5) (2009) 056117.
- [238] Z. Yang, R. Algesheimer, C.J. Tessone, A comparative analysis of community detection algorithms on artificial networks, *Sci. Rep.* 6 (2016) 30750.
- [239] V.D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks, *J. Stat. Mech. Theory Exp.* 2008 (10) (2008) P10008.
- [240] S. Sobolevsky, R. Campari, A. Belyi, C. Ratti, General optimization technique for high-quality community detection in complex networks, *Phys. Rev. E* 90 (1) (2014) 012811.
- [241] S. Fortunato, M. Barthélémy, Resolution limit in community detection, *Proc. Natl. Acad. Sci.* 104 (1) (2007) 36–41.
- [242] A. Lancichinetti, F. Radicchi, J.J. Ramasco, S. Fortunato, Finding statistically significant communities in networks, *PLoS One* 6 (4) (2011) e18961.
- [243] P.J. Mucha, T. Richardson, K. Macon, M.A. Porter, J.-P. Onnela, Community structure in time-dependent, multiscale, and multiplex networks, *Science* 328 (5980) (2010) 876–878.
- [244] S.L. Pimm, J.H. Lawton, Are food webs divided into compartments? *J. Anim. Ecol.* (1980) 879–898.
- [245] D. Raffaelli, S.J. Hall, Compartments and predation in an estuarine food web, *J. Anim. Ecol.* (1992) 551–560.
- [246] L. Dicks, S. Corbet, R. Pywell, Compartmentalization in plant-insect flower visitor webs, *J. Anim. Ecol.* 71 (1) (2002) 32–43.
- [247] D.B. Stouffer, J. Bascompte, Compartmentalization increases food-web persistence, *Proc. Natl. Acad. Sci.* 108 (9) (2011) 3648–3652.
- [248] J. Grilli, T. Rogers, S. Allesina, Modularity and stability in ecological communities, *Nature Commun.* 7 (2016) 12031.
- [249] L.J. Gilarranz, B. Rayfield, G. Liñán-Cembrano, J. Bascompte, A. González, Effects of network modularity on the spread of perturbation impact in experimental metapopulations, *Science* 357 (6347) (2017) 199–201.
- [250] R.H. Whittaker, Vegetation of the Siskiyou mountains, Oregon and California, *Ecol. Monograph* 30 (3) (1960) 279–338.
- [251] M. Mitchell, Complex systems: Network thinking, *Artificial Intelligence* 170 (18) (2006) 1194–1212.
- [252] M. Mitchell, Complexity: A Guided Tour, Oxford University Press, 2009.
- [253] M.E. Newman, Complex systems: A survey, 2011, arXiv preprint arXiv:1112.1440.
- [254] M. Medo, G. Cimini, S. Gualdi, Temporal effects in the growth of networks, *Phys. Rev. Lett.* 107 (23) (2011) 238701.
- [255] M. Medo, Statistical validation of high-dimensional models of growing networks, *Phys. Rev. E* 89 (3) (2014) 032801.
- [256] L. Santamaría, M.A. Rodríguez-Gironés, Linkage rules for plant-pollinator networks: trait complementarity or exploitation barriers? *PLoS Biol.* 5 (2) (2007) e31.
- [257] E.L. Rezende, P. Jordano, J. Bascompte, Effects of phenotypic complementarity and phylogeny on the nested structure of mutualistic networks, *Oikos* 116 (11) (2007) 1919–1929.
- [258] K. Trøjelsgaard, J.M. Olesen, Macroecology of pollination networks, *Global Ecol. Biogeogr.* 22 (2) (2013) 149–162.
- [259] M. Schleuning, L. Ingmann, R. Strauss, S.A. Fritz, B. Dalsgaard, D. Matthias Dehling, M. Plein, F. Saavedra, B. Sandel, J.-C. Svenning, et al., Ecological, historical and evolutionary determinants of modularity in weighted seed-dispersal networks, *Ecol. Lett.* 17 (4) (2014) 454–463.
- [260] E.A. Welti, A. Joern, Structure of trophic and mutualistic networks across broad environmental gradients, *Ecol. Evol.* 5 (2) (2015) 326–334.
- [261] B. Dalsgaard, K. Trøjelsgaard, A.M. Martín González, D. Nogués-Bravo, J. Ollerton, T. Petanidou, B. Sandel, M. Schleuning, Z. Wang, C. Rahbek, et al., Historical climate-change influences modularity and nestedness of pollination networks, *Ecography* 36 (12) (2013) 1331–1340.
- [262] E. Sebastián-González, B. Dalsgaard, B. Sandel, P.R. Guimaraes, Macroecological trends in nestedness and modularity of seed-dispersal networks: human impact matters, *Global Ecol. Biogeogr.* 24 (3) (2015) 293–303.
- [263] K. Takemoto, K. Kajihara, Human impacts and climate change influence nestedness and modularity in food-web and mutualistic networks, *PLoS One* 11 (6) (2016) e0157929.
- [264] L. Pellissier, C. Albouy, J. Bascompte, N. Farwig, C. Graham, M. Loreau, M.A. Magliaresi, C.J. Melián, C. Pitteloud, T. Roslin, R. Rohr, S. Saavedra, W. Thuiller, G. Woodward, N.E. Zimmermann, D. Gravel, Comparing species interaction networks along environmental gradients, *Biol. Rev.* 93 (2) (2018) 785–800.
- [265] H. Tuomisto, A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity, *Ecography* 33 (1) (2010) 2–22.
- [266] T.A. Sørensen, A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons, *K. Dan. Vidensk. Selsk. Biol. Skr.* 5 (1948) 1–34.
- [267] A. Baselga, Partitioning the turnover and nestedness components of beta diversity, *Global Ecol. Biogeogr.* 19 (1) (2010) 134–143.

- [268] G.G. Simpson, Mammals and the nature of continents, *Am. J. Sci.* 241 (1) (1943) 1–31.
- [269] A. Baselga, The relationship between species replacement, dissimilarity derived from nestedness, and nestedness, *Global Ecol. Biogeogr.* 21 (12) (2012) 1223–1232.
- [270] A. Baselga, C.D.L. Orme, Betapart: an r package for the study of beta diversity, *Methods Ecol. Evol.* 3 (5) (2012) 808–812.
- [271] S. Azalee, S. Suweis, J. Grilli, I. Volkov, J.R. Banavar, A. Maritan, Statistical mechanics of ecological systems: Neutral theory and beyond, *Rev. Modern Phys.* 88 (3) (2016) 035003.
- [272] J. Rosindell, S.P. Hubbell, R.S. Etienne, The unified neutral theory of biodiversity and biogeography at age ten, *Trends Ecol. Evol.* 26 (7) (2011) 340–348.
- [273] R.S. Etienne, A new sampling formula for neutral biodiversity, *Ecol. Lett.* 8 (3) (2005) 253–260.
- [274] J.M. Olesen, J. Bascompte, Y.L. Dupont, H. Elberling, C. Rasmussen, P. Jordano, Missing and forbidden links in mutualistic networks, *Proc. R. Soc. Lond. Biol.* (2010) rspb20101371.
- [275] E. Canard, N. Mouquet, L. Marescot, K.J. Gaston, D. Gravel, D. Mouillot, Emergence of structural patterns in neutral trophic networks, *PLoS One* 7 (8) (2012) e38295.
- [276] J. Ollerton, S.D. Johnson, L. Cranmer, S. Kellie, The pollination ecology of an assemblage of grassland asclepiads in South Africa, *Ann. Botany* 92 (6) (2003) 807–834.
- [277] P. Jordano, J. Bascompte, J.M. Olesen, Invariant properties in coevolutionary networks of plant–animal interactions, *Ecol. Lett.* 6 (1) (2003) 69–81.
- [278] E.L. Rezende, J.E. Lavabre, P.R. Guimarães, P. Jordano, J. Bascompte, Non-random coextinctions in phylogenetically structured mutualistic networks, *Nature* 448 (7156) (2007) 925.
- [279] D.P. Vázquez, N.P. Chacoff, L. Cagnolo, Evaluating multiple determinants of the structure of plant–animal mutualistic networks, *Ecology* 90 (8) (2009) 2039–2046.
- [280] Z.-M. Ren, A. Zeng, Y.-C. Zhang, Structure-oriented prediction in complex networks, *Phys. Rep.* 750 (2018) 1–51.
- [281] M. Baldovin, F. Cecconi, M. Cencini, A. Puglisi, A. Vulpiani, The role of data in model building and prediction: A survey through examples, *Entropy* 20 (10) (2018) 807.
- [282] L. Lü, T. Zhou, Link prediction in complex networks: A survey, *Physica A* 390 (6) (2011) 1150–1170.
- [283] T. Squartini, G. Caldarelli, G. Cimini, A. Gabrielli, D. Garlaschelli, Reconstruction methods for networks: the case of economic and financial systems, *Phys. Rep.* 757 (2018) 1–47.
- [284] M. Maron, R. Mac Nally, D. M Watson, A. Lill, Can the biotic nestedness matrix be used predictively? *Oikos* 106 (3) (2004) 433–444.
- [285] M. Medo, M. Mariani, L. Lü, Link prediction in bipartite nested networks, *Entropy* 20 (10) (2018) 777.
- [286] C.A. Hidalgo, B. Klinger, A.-L. Barabási, R. Hausmann, The product space conditions the development of nations, *Science* 317 (5837) (2007) 482–487.
- [287] A. Zaccaria, M. Cristelli, A. Tacchella, L. Pietronero, How the taxonomy of products drives the economic development of countries, *PLoS One* 9 (12) (2014) e113770.
- [288] A. Vidmer, A. Zeng, M. Medo, Y.-C. Zhang, Prediction in complex systems: The case of the international trade network, *Physica A* 436 (2015) 188–199.
- [289] E.N. Lorenz, Atmospheric predictability as revealed by naturally occurring analogues, *J. Atmos. Sci.* 26 (4) (1969) 636–646.
- [290] R.V. Noorden, Physicists make ‘weather forecasts’ for economies, *Nature* 1038 (2015) 16963.
- [291] H. Liao, M.S. Mariani, M. Medo, Y.-C. Zhang, M.-Y. Zhou, Ranking in evolving complex networks, *Phys. Rep.* 689 (2017) 1–54.
- [292] E. Pugliese, G.L. Chiarotti, A. Zaccaria, L. Pietronero, Complex economies have a lateral escape from the poverty trap, *PLoS One* 12 (1) (2017) e0168540.
- [293] G. Vinci, R. Benzi, Economic complexity: Correlations between gross domestic product and fitness, *Entropy* 20 (10) (2018) 766.
- [294] O. Angelini, M. Cristelli, A. Zaccaria, L. Pietronero, The complex dynamics of products and its asymptotic properties, *PLoS One* 12 (5) (2017) e0177360.
- [295] R. Hausmann, C.A. Hidalgo, S. Bustos, M. Coscia, A. Simoes, M.A. Yildirim, *The Atlas of Economic Complexity: mapping Paths to Prosperity*, Mit Press, 2014.
- [296] A. Tacchella, R. Di Clemente, A. Gabrielli, L. Pietronero, The build-up of diversity in complex ecosystems, 2016, arXiv preprint arXiv:1609.03617.
- [297] S.N. Dorogovtsev, J.F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW*, OUP Oxford, 2013.
- [298] G. Caldarelli, A. Capocci, P. De Los Rios, M.A. Munoz, Scale-free networks from varying vertex intrinsic fitness, *Phys. Rev. Lett.* 89 (25) (2002) 258702.
- [299] F. Zhang, C. Hui, J.S. Terblanche, An interaction switch predicts the nested architecture of mutualistic networks, *Ecol. Lett.* 14 (8) (2011) 797–803.
- [300] M. Boguná, R. Pastor-Satorras, Class of correlated random networks with hidden variables, *Phys. Rev. E* 68 (3) (2003) 036112.
- [301] M. Granovetter, Threshold models of collective behavior, *Am. J. Sociol.* 83 (6) (1978) 1420–1443.
- [302] D.J. Watts, A simple model of global cascades on random networks, *Proc. Natl. Acad. Sci.* 99 (9) (2002) 5766–5771.
- [303] D. Kempe, J. Kleinberg, É. Tardos, Maximizing the spread of influence through a social network, in: *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, 2003, pp. 137–146.
- [304] S. Valverde, The spandrels of ecological architecture, 2018, <https://natureecocommunity.nature.com/users/68857-sergi-valverde/posts/23428-the-spandrels-of-ecological-architecture>, [Online; Accessed 26 August 2018].
- [305] E. Burgos, H. Ceva, R.P. Perazzo, M. Devoto, D. Medan, M. Zimmermann, A.M. Delbue, Why nestedness in mutualistic networks? *J. Theoret. Biol.* 249 (2) (2007) 307–313.
- [306] R.P. Perazzo, L. Hernández, H. Ceva, E. Burgos, J.I. Alvarez-Hamelin, Study of the influence of the phylogenetic distance on the interaction network of mutualistic ecosystems, *Physica A* 394 (2014) 124–135.
- [307] P.R. Guimaraes Jr, G. Machado, M.A. de Aguiar, P. Jordano, J. Bascompte, A. Pinheiro, S.F. Dos Reis, Build-up mechanisms determining the topology of mutualistic networks, *J. Theoret. Biol.* 249 (2) (2007) 181–189.
- [308] K. Takemoto, M. Arita, Nested structure acquired through simple evolutionary process, *J. Theoret. Biol.* 264 (3) (2010) 782–786.
- [309] D. Schöch, T.W. Valente, U. Brandes, Correlations among centrality indices and a class of uniquely ranked graphs, *Social Networks* 50 (2017) 46–54.
- [310] A.-L. Barabási, R. Albert, H. Jeong, Mean-field theory for scale-free random networks, *Physica A* 272 (1–2) (1999) 173–187.
- [311] C.F. Manski, The structure of random utility models, *Theory and Decision* 8 (3) (1977) 229–254.
- [312] C. Ballester, A. Calvó-Armengol, Y. Zenou, Who’s who in networks. Wanted: The key player, *Econometrica* 74 (5) (2006) 1403–1417.
- [313] A. Calvó-Armengol, E. Patacchini, Y. Zenou, Peer effects and social networks in education, *Rev. Econom. Stud.* 76 (4) (2009) 1239–1267.
- [314] T.H. Feeley, G.A. Barnett, Predicting employee turnover from communication networks, *Hum. Commun. Res.* 23 (3) (1997) 370–387.
- [315] T.H. Feeley, S.-I. Moon, R.S. Kozey, A.S. Slowe, An erosion model of employee turnover based on network centrality, *J. Appl. Commun. Res.* 38 (2) (2010) 167–188.

- [316] J. Yuan, Q.-M. Zhang, J. Gao, L. Zhang, X.-S. Wan, X.-J. Yu, T. Zhou, Promotion and resignation in employee networks, *Physica A* 444 (2016) 442–447.
- [317] N.E. Friedkin, Theoretical foundations for centrality measures, *Am. J. Sociol.* 96 (6) (1991) 1478–1504.
- [318] D. Taylor, D.B. Larremore, Social climber attachment in forming networks produces a phase transition in a measure of connectivity, *Phys. Rev. E* 86 (3) (2012) 031140.
- [319] P.L. Krapivsky, S. Redner, Organization of growing random networks, *Phys. Rev. E* 63 (6) (2001) 066123.
- [320] S. Saavedra, R.P. Rohr, J.M. Olesen, J. Bascompte, Nested species interactions promote feasibility over stability during the assembly of a pollinator community, *Ecol. Evol.* 6 (4) (2016) 997–1007.
- [321] S.V. Robinson, G. Losapio, G.H. Henry, Flower-power: Flower diversity is a stronger predictor of network structure than insect diversity in an Arctic plant-pollinator network, *Ecol. Complex.* 36 (2018) 1–6.
- [322] W.J. Sutherland, R.P. Freckleton, H.C.J. Godfray, S.R. Beissinger, T. Benton, D.D. Cameron, Y. Carmel, D.A. Coomes, T. Coulson, M.C. Emmerson, et al., Identification of 100 fundamental ecological questions, *J. Ecol.* 101 (1) (2013) 58–67.
- [323] L.P. Koh, R.R. Dunn, N.S. Sodhi, R.K. Colwell, H.C. Proctor, V.S. Smith, Species coextinctions and the biodiversity crisis, *Science* 305 (5690) (2004) 1632–1634.
- [324] E. Thébault, C. Fontaine, Stability of ecological communities and the architecture of mutualistic and trophic networks, *Science* 329 (5993) (2010) 853–856.
- [325] A. James, J.W. Pitchford, M.J. Plank, Disentangling nestedness from models of ecological complexity, *Nature* 487 (7406) (2012) 227.
- [326] S. Allesina, S. Tang, The stability-complexity relationship at age 40: a random matrix perspective, *Popul. Ecol.* 57 (1) (2015) 63–75.
- [327] C.N. Kaiser-Bunbury, N. Blüthgen, Integrating network ecology with applied conservation: a synthesis and guide to implementation, *AoB Plants* 7 (2015).
- [328] S. Battiston, M. Puliga, R. Kaushik, P. Tasca, G. Caldarelli, Debtrank: Too central to fail? financial networks, the fed and systemic risk, *Sci. Rep.* 2 (2012) 541.
- [329] S. Battiston, J.D. Farmer, A. Flache, D. Garlaschelli, A.G. Haldane, H. Heesterbeek, C. Hommes, C. Jaeger, R. May, M. Scheffer, Complexity theory and financial regulation, *Science* 351 (6275) (2016) 818–819.
- [330] J. Memmott, N.M. Waser, M.V. Price, Tolerance of pollination networks to species extinctions, *Proc. R. Soc. Lond. Biol.* 271 (1557) (2004) 2605–2611.
- [331] M.A. Aizen, M. Sabatino, J.M. Tylianakis, Specialization and rarity predict nonrandom loss of interactions from mutualist networks, *Science* 335 (6075) (2012) 1486–1489.
- [332] E. Burgos, H. Ceva, L. Hernández, R. Perazzo, Understanding and characterizing nestedness in mutualistic bipartite networks, *Comput. Phys. Comm.* 180 (4) (2009) 532–535.
- [333] F.E. Clements, F.L. Long, Experimental Pollination: An Outline of the Ecology of Flowers and Insects, no. 336, Carnegie Institution of Washington, 1923.
- [334] C.N. Kaiser-Bunbury, S. Muff, J. Memmott, C.B. Müller, A. Caflisch, The robustness of pollination networks to the loss of species and interactions: a quantitative approach incorporating pollinator behaviour, *Ecol. Lett.* 13 (4) (2010) 442–452.
- [335] S. Allesina, M. Pascual, Googling food webs: can an eigenvector measure species' importance for coextinctions? *PLoS Comput. Biol.* 5 (9) (2009) e1000494.
- [336] J. Gao, B. Barzel, A.-L. Barabási, Universal resilience patterns in complex networks, *Nature* 530 (7590) (2016) 307.
- [337] T. Okuyama, J.N. Holland, Network structural properties mediate the stability of mutualistic communities, *Ecol. Lett.* 11 (3) (2008) 208–216.
- [338] S. Saavedra, D.B. Stouffer, "Disentangling nestedness" disentangled, *Nature* 500 (7463) (2013) E1.
- [339] A. James, J.W. Pitchford, M.J. Plank, James et al. reply, *Nature* 500 (7463) (2013) E2.
- [340] J. Bascompte, A. Ferrera, A structural theory of mutualistic networks, in: *Theoretical Ecology: Principles and Applications*, Kevin S. McCann and Gabriel Gellner, Editors. Oxford University Press, in revision.
- [341] A. Vulpiani, *Chaos: From Simple Models to Complex Systems*, vol. 17, World Scientific, 2010.
- [342] A. Lotka, Zur theorie der periodischen reaktionen, *Z. Phys. Chem.* 72 (1) (1910) 508–511.
- [343] V. Volterra, Fluctuations in the abundance of a species considered mathematically, *Nature* 118 (1926) 558–560.
- [344] J.D. Murray, *Mathematical Biology I. An Introduction*, Springer, 2002.
- [345] J. Grilli, M. Adorisio, S. Suweis, G. Barabás, J.R. Banavar, S. Allesina, A. Maritan, Feasibility and coexistence of large ecological communities, *Nature Commun.* 8 (2017).
- [346] L. Stone, The feasibility and stability of large complex biological networks: a random matrix approach, *Sci. Reports* 8 (1) (2018) 8246.
- [347] R. Levins, *Evolution in Changing Environments: Some Theoretical Explorations*, no. 2, Princeton University Press, 1968.
- [348] S. Saavedra, R.P. Rohr, J. Bascompte, O. Godoy, N.J. Kraft, J.M. Levine, A structural approach for understanding multispecies coexistence, *Ecol. Monograph* 87 (3) (2017) 470.
- [349] B. Goh, Stability in models of mutualism, *Amer. Nat.* 113 (2) (1979) 261–275.
- [350] U. Bastolla, M. Lässig, S.C. Manrubia, A. Valleriani, Biodiversity in model ecosystems, I: coexistence conditions for competing species, *J. Theoret. Biol.* 235 (4) (2005) 521–530.
- [351] U. Bastolla, M. Lässig, S.C. Manrubia, A. Valleriani, Biodiversity in model ecosystems, II: species assembly and food web structure, *J. Theoret. Biol.* 235 (4) (2005) 531–539.
- [352] D.H. Wright, A simple, stable model of mutualism incorporating handling time, *Amer. Nat.* 134 (4) (1989) 664–667.
- [353] C.S. Holling, Some characteristics of simple types of predation and parasitism, *Canad. Entomol.* 91 (7) (1959) 385–398.
- [354] J. Bascompte, P. Jordano, J.M. Olesen, Asymmetric coevolutionary networks facilitate biodiversity maintenance, *Science* 312 (5772) (2006) 431–433.
- [355] J.N. Holland, T. Okuyama, D.L. DeAngelis, Comment on "asymmetric coevolutionary networks facilitate biodiversity maintenance", *Science* 313 (5795) (2006) 1887–1887.
- [356] R. May, A.R. McLean, *Theoretical Ecology: Principles and Applications*, Blackwell Scientific Publications, Oxford, 1981.
- [357] S. Saavedra, R.P. Rohr, V. Dakos, J. Bascompte, Estimating the tolerance of species to the effects of global environmental change, *Nature Commun.* 4 (2013) 2350.
- [358] D.P. Vázquez, C.J. Melián, N.M. Williams, N. Blüthgen, B.R. Krasnov, R. Poulin, Species abundance and asymmetric interaction strength in ecological networks, *Oikos* 116 (7) (2007) 1120–1127.
- [359] C. Gracia-Lázaro, L. Hernández, J. Borge-Holthoefer, Y. Moreno, The joint influence of competition and mutualism on the biodiversity of mutualistic ecosystems, *Sci. Reports* 8 (1) (2018) 9253.
- [360] G. Losapio, M. Fortuna, J. Bascompte, B. Schmid, R. Michalet, R. Neumeyer, L. Castro, P. Cerretti, C. Germann, J.-P. Haenni, et al., Facilitation between plants shapes pollination networks, *bioRxiv* (2017) 161034.
- [361] F.S. Valdovinos, P.M. De Espanés, J.D. Flores, R. Ramos-Jiliberto, Adaptive foraging allows the maintenance of biodiversity of pollination networks, *Oikos* 122 (6) (2013) 907–917.

- [362] F.S. Valdovinos, B.J. Brosi, H.M. Briggs, P. Moisset de Espanés, R. Ramos-Jiliberto, N.D. Martinez, Niche partitioning due to adaptive foraging reverses effects of nestedness and connectance on pollination network stability, *Ecol. Lett.* 19 (10) (2016) 1277–1286.
- [363] F. Morone, G. Del Ferraro, H. Makse, The k-core as a predictor of structural collapse in mutualistic ecosystems, *Nat. Phys.* (2018).
- [364] P. Alberch, The logic of monsters: evidence for internal constraint in development and evolution, *Geobios* 22 (1989) 21–57.
- [365] D.D. Šiljak, Large-Scale Dynamic Systems: Stability and Structure, vol. 2, North Holland, 1978.
- [366] J.H. Vandermeer, Interspecific competition: a new approach to the classical theory, *Science* 188 (4185) (1975) 253–255.
- [367] R.V. Solé, J. Valls, On structural stability and chaos in biological systems, *J. Theoret. Biol.* 155 (1) (1992) 87–102.
- [368] T.J. Case, An Illustrated Guide to Theoretical Ecology, Oxford University Press, 2000.
- [369] C. Song, R.P. Rohr, S. Saavedra, A guideline to study the feasibility domain of multi-trophic and changing ecological communities, *J. Theoret. Biol.* 450 (2018) 30–36.
- [370] J.M. Ribando, Measuring solid angles beyond dimension three, *Discrete Comput. Geom.* 36 (3) (2006) 479–487.
- [371] S. Cencí, S. Saavedra, Structural stability of nonlinear population dynamics, *Phys. Rev. E* 97 (1) (2018) 012401.
- [372] J.M. Olesen, J. Bascompte, H. Elberling, P. Jordano, Temporal dynamics in a pollination network, *Ecology* 89 (6) (2008) 1573–1582.
- [373] V.L. Girko, Circular law, *Theory Probab. Appl.* 29 (4) (1985) 694–706.
- [374] T. Tao, V. Vu, Random matrices: Universality of ESDs and the circular law, *Ann. Probab.* (2010) 2023–2065.
- [375] S.L. Pimm, Complexity and stability: another look at MacArthur's original hypothesis, *Oikos* (1979) 351–357.
- [376] L. Stone, The google matrix controls the stability of structured ecological and biological networks, *Nature Commun.* 7 (2016) 12857.
- [377] S. Suweis, J. Grilli, J.R. Banavar, S. Allesina, A. Maritan, Effect of localization on the stability of mutualistic ecological networks, *Nature Commun.* 6 (2015) 10179.
- [378] S. Saavedra, R.P. Rohr, L.J. Gilarranz, J. Bascompte, How structurally stable are global socioeconomic systems? *J. R. Soc. Interface* 11 (100) (2014) 20140693.
- [379] S. Fattorini, Non-randomness in the species-area relationship: testing the underlying mechanisms, *Oikos* 116 (4) (2007) 678–689.
- [380] S.J. Beckett, H.T. Williams, Coevolutionary diversification creates nested-modular structure in phage–bacteria interaction networks, *Interface Focus* 3 (6) (2013) 20130033.
- [381] G. Fagiolo, The international-trade network: gravity equations and topological properties, *J. Econ. Interact. Coord.* 5 (1) (2010) 1–25.
- [382] P. Expert, T.S. Evans, V.D. Blondel, R. Lambiotte, Uncovering space-independent communities in spatial networks, *Proc. Natl. Acad. Sci.* 108 (19) (2011) 7663–7668.
- [383] C.O. Flores, S. Valverde, J.S. Weitz, Multi-scale structure and geographic drivers of cross-infection within marine bacteria and phages, *ISME J.* 7 (3) (2013) 520–532.
- [384] A. Grimm, C.J. Tessone, Detecting nestedness in graphs, in: International Workshop on Complex Networks and their Applications, Springer, 2016, pp. 171–182.
- [385] S. Kojaku, N. Masuda, Finding multiple core-periphery pairs in networks, *Phys. Rev. E* 96 (2017) 052313.
- [386] M. Palazzi, J. Borge-Holthoefer, C. Tessone, A. Solé-Ribalta, Antagonistic structural patterns in complex networks, 2018, arXiv preprint arXiv:1810.12785.
- [387] S. Kojaku, M. Xu, H. Xia, N. Masuda, Multiscale core-periphery structure in a global liner shipping network, *Sci. Reports* 9 (1) (2019) 404.
- [388] B.W. Kernighan, S. Lin, An efficient heuristic procedure for partitioning graphs, *Bell Syst. Tech. J.* 49 (2) (1970) 291–307.
- [389] M.P. Van Den Heuvel, O. Sporns, Rich-club organization of the human connectome, *J. Neurosci.* 31 (44) (2011) 15775–15786.
- [390] M. De Domenico, A. Solé-Ribalta, E. Omodei, S. Gómez, A. Arenas, Ranking in interconnected multilayer networks reveals versatile nodes, *Nature Commun.* 6 (2015) 6868.
- [391] A. Solé-Ribalta, C.J. Tessone, C.G. Ferrari, J. Borge-Holthoefer, Disentangling co-occurrence patterns in n-partite ecosystems, 2018, arXiv preprint arXiv:1807.04666.
- [392] C. Fontaine, P.R. Guimarães Jr., S. Kéfi, N. Loeuille, J. Memmott, W.H. van Der Putten, F.J. van Veen, E. Thébaud, The ecological and evolutionary implications of merging different types of networks, *Ecol. Lett.* 14 (11) (2011) 1170–1181.
- [393] S. Suweis, J. Grilli, A. Maritan, Disentangling the effect of hybrid interactions and of the constant effort hypothesis on ecological community stability, *Oikos* 123 (5) (2014) 525–532.
- [394] S. Pilosof, M.A. Porter, M. Pascual, S. Kéfi, The multilayer nature of ecological networks, *Nat. Ecol. Evol.* 1 (4) (2017) 0101.
- [395] M. Barigozzi, G. Fagiolo, D. Garlaschelli, Multinetwork of international trade: A commodity-specific analysis, *Phys. Rev. E* 81 (4) (2010) 046104.
- [396] G. Strona, P. Galli, D. Seveso, S. Montano, S. Fattorini, et al., Nestedness for dummies (ned): A user-friendly web interface for exploratory nestedness analysis, *J. Stat. Softw.* 59 (c03) (2014).
- [397] L. Lü, D. Chen, X.-L. Ren, Q.-M. Zhang, Y.-C. Zhang, T. Zhou, Vital nodes identification in complex networks, *Phys. Rep.* 650 (2016) 1–63.
- [398] S.P. Borgatti, D.S. Halgin, Analyzing affiliation networks, *Sage Handb. Soc. Netw. Anal.* 1 (2011) 417–433.
- [399] P. Bonacich, Factoring and weighting approaches to status scores and clique identification, *J. Math. Sociol.* 2 (1) (1972) 113–120.
- [400] S. Brin, L. Page, The anatomy of a large-scale hypertextual web search engine, *Comput. Netw. SDN Syst.* 30 (1) (1998) 107–117.