Structural controllability of ecological networks

1   **E. Fernando Cagua[1], Kate L. Wootton[1,2], Daniel B. Stouffer[1]**

2   [1] Centre for Integrative Ecology, School of Biological Sciences, University of Canterbury, Private Bag

3   4800, Christchurch 8041, New Zealand

4   [2] Current address: Department of Ecology, Swedish University of Agricultural Sciences, Box 7044, SE-750

5   07 Uppsala, Sweden

6   **Author for correspondence:** Daniel Stouffer (daniel.stouffer@canterbury.ac.nz) - +64 3 364 2729 -

7   Centre for Integrative Ecology, School of Biological Sciences, University of Canterbury, Private Bag 4800,

8   Christchurch 8140, New Zealand

9   *Alternative titles:*

10 • Keystoneness, centrality, and the structural controllability of ecological networks

Introduction

A major goal in ecology is to understand the roles played by different species in the biotic environment. Within community ecology, a complex-systems approach has led to the development of a variety analytical and simulation tools with which to compare and contrast the roles of species embedded in a network of interactions (J. Bascompte & Stouffer, 2009; Coux, Rader, Bartomeus, & Tylianakis, 2016; Guimerà & Amaral, 2005; Stouffer, Sales-Pardo, Sirer, & Bascompte, 2012). A particularly relevant dimension of any species' role is its ability to alter the abundance of other species in the community—since changes of this nature impact ecosystem function and biodiversity, and might have knock-on effects on other processes and services. This ability is often referred to as a species' "keystoneness" (Mills & Doak, 1993).

A significant proportion of the network tools used to estimate species roles in this context rely on the calculation of a species' centrality—a relative ranking of its importance that stems from social network research (Friedkin, 1991). Generally speaking, central species tend to be present in more food chains, are better connected, and have broader niches. Centrality metrics have been shown to be useful tools to rank species in regards to their risk of extinction or the potential secondary extinctions after the species is lost (Dunne, Williams, & Martinez, 2002; Kaiser-Bunbury, Muff, Memmott, Müller, & Caflisch, 2010). Although there are some differences among centrality metrics—as each of them emphasizes a different aspect of a species' interaction patterns—they all tend to be correlated with the way species participate in the network's food chains (Jordán, Benedek, & Podani, 2007). Because of this property, centrality metrics have been often used to identify keystone species.

This relationship between centrality and keystoneness relies on an important assumption. That that there is a link between a given species' centrality and its ability to alter the abundances of other species. The basis of this assumption is reasonably clear: a major justification for representing ecological communities as networks is the idea that the abundance of a species depends on the abundance of its neighbours (those species with which it interacts) and, by extension, the abundance of its neighbours' neighbours, and so on. By definition, more central species are better connected and hence should be more likely to affect the abundance of other species in the network.

Despite being conceptually intuitive, the relationship between centrality and keystoneness is largely phenomenological. Often, dramatic changes in ecosystem functioning can also occur without complete removal of a species. Hence, we expect the usefulness of centrality to diminish as we move beyond species removal towards understanding how species relate to the state and stability of the ecosystem. Particularly so, when the state of an ecosystem is underpinned by more than a single species. As a result, community ecology could arguably benefit from an alternative more mechanistically-grounded approach to understand how species affect each other's abundance.

Species abundances (and consequently the state of the community) are influenced by both their interactions, the specific dynamics of these interactions and the mechanisms of self-regulation. However, community and population dynamics can be modelled in innumerable ways, and empirical support for one versus another is often still ambiguous. The alternative approach should, therefore, acknowledge ecosystem dynamics, but without being overly dependent on the particular choices of how they are characterised. Among the various possibilities, *control theory* appears to be a strong candidate (Isbell & Loreau, 2013). Control theory is a widely-studied branch of engineering used to determine and supervise the behaviour of dynamical systems (Motter, 2015). It is inherently designed to deal with system feedbacks and its application has recently been expanded to complex networks (Liu & Barabási, 2016). In concordance with long-standing ecological questions, advances in this field have established a clear link between the structure of the network and the way nodes affect each other. Unlike centrality indices, however, this link is not based on a priori assumptions between network metrics and keystoneness but is instead based on well-established advances in both dynamical and complex-systems theory.

At its fundamental level, control theory first determines whether a system is controllable or not; that is, it asks if a system can be driven to a desired state within a finite amount of time. Although the controllability of a network is a whole-system property, it has been recently shown that asking for the controllability of a complex-system is equivalent to finding a particular set of nodes: the set with which is possible to control the state of the whole network. Notably, control theory is consistent with the observation that to change the state of a network, more than a ranking, we need a set of nodes. Importantly, this set of nodes is not always unique for a given network. This implies that an examination of the distinct sets provides a means to connect nodes with its *general* ability to modify the system to which they belong.

Here, we apply methods from control theory to a particular ecological problem and show how it can be used to gain insight into the role of species in an ecological network. Specifically, we outline the approach using a set of ten pairs of uninvaded and invaded plant-pollinator communities. We use invaded communities because there is strong empirical evidence showing that invasive species play an important role shaping the abundances of other species, something which is particularly true in these ten networks (Bartomeus, Vilà, & Santamaría, 2008; Lopezaraiza–Mikel, Hayes, Whalley, & Memmott, 2007). This choice thus offers us an opportunity to explicitly contrast our theoretical observations with empirical evidence. With these networks, we first ask whether there are differences between the controllability of invaded and uninvaded networks. This question is motivated by the difficulties commonly encountered with invasive-species eradication and ecosystem restoration (Woodford et al., 2016). We then expand existing methods from control theory to effectively link the *controllability* of a network with the role of particular species. We ask—from a control-theoretic perspective—whether there are key differences between species in the role they play at driving the population of other members of the community and

identify the ecological factors related to these differences. This allows us to identify species that are

78 critical to changes in the ecosystem state and show that they have a larger than expected impact on the

79 structural stability of the community. Finally, we compare the proposed approach to current methods

80 based on species centrality and show how these methods are indeed valuable but ultimately paint a

81 limited picture in regards to the dynamic "keystoneness" of a species.

Methods

82 We used ten paired pollination communities to apply the control-theoretic approach. Each community

83 pair was composed of a community invaded by a plant and a community free of the invasive species.

84 Four pairs correspond to natural or semi-natural vegetation communities in the city of Bristol, UK

85 (Lopezaraiza–Mikel et al., 2007). These communities are comprised of 19–87 species (mean 55), and

86 non-invaded plots were obtained by experimentally removing all the flowers of the invasive species

87 *Impatients grandulifera*. The other six pairs were obtained from lower diversity Mediterranean shrublands

88 in Cap de Creus National Park, Spain (Bartomeus et al., 2008). These communities are comprised of

89 30–57 species (mean 38); in contrast to the above, uninvaded communities were obtained from plots

90 that had not yet been colonised by either of the invasive species *Carpobrotus affine acinaciformis* or

91 *Opuntia stricta*. The structure of all these communities was defined by the pollinator visitation frequency,

92 which has been shown to be an appropriate surrogate for interspecific effects in pollination networks (J.

93 Bascompte, Jordano, & Olesen, 2006; Diego P. Vázquez, Morris, & Jordano, 2005). Full details about the

94 empirical networks can be found in the Supporting Information Section S1.

95 The first step in applying methods of control theory is to construct a directed network that is able to

96 provide an indication of the extent to which species affect each other's abundance. In some ecological

97 networks, establishing the directionality can be relatively straightforward, for example when links represent

98 biomass transfer or energy flow (Isbell & Loreau, 2013). In pollination networks, however, this directionality

99 is less obvious as both species can, in theory, benefit from the interaction. We overcome that obstacle by

100 noting that the extent to which species i affects species j relative to the extent to which j affects i can be

101 summarised by the interaction asymmetry (J. Bascompte et al., 2006). This asymmetry is given by

$$a(i,j) = a(j,i) = \frac{|d_{ij} - d_{ji}|}{max\,(d_{ij}, d_{ji})},$$

102 where the dependence of plant i on pollinator j, d_{ij}, is the proportion of the visits from pollinator j

103 compared to all pollinator visits to plant i. Previous research has shown that mutualistic interactions are

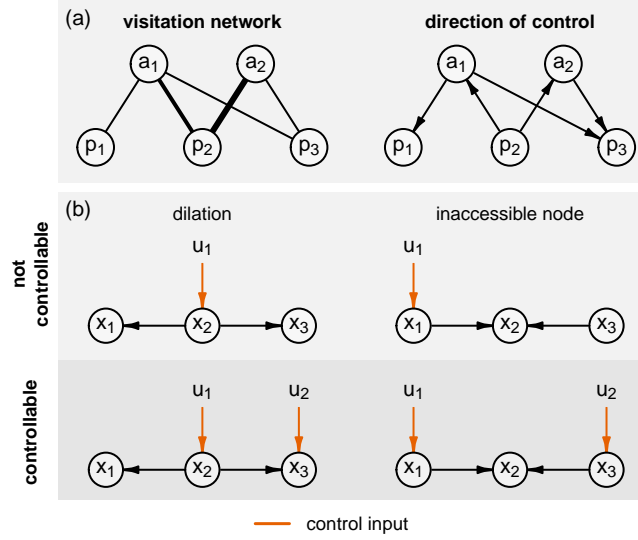104 often highly asymmetric in natural communities; in other words, if a plant species is largely dependent

Figure 1: Direction of control and controllability conditions. (a) To establish the direction of control we start with a weighted visitation network (on the left). In this network, the width of the links corresponds to the frequency of visitation between animals a_i and plants p_i, with wider links indicating more visits. Plant p_1 is visited exclusively by a_1 but p_1 represents only a small fraction of the floral resources of a_1. Therefore the population of p_1 is more likely to be affected by a_1 than vice-versa. We represent this with a directed link from a_1 to p_1 (right). The direction of control between all other species pairs can be similarly determined by inspecting the difference between their relative dependences. (b) Once we have established the directions of control we can determine whether the network is controllable or not. A system defined by a directed network (with state nodes x_i; species populations in an ecological context) and external control inputs (nodes u_i, orange links) is structurally controllable if it satisfies two conditions: it has no dilations (expansions in the network) and no inaccessible nodes. The system on the top left is not controllable because there is a dilation in which node x_2 is being used to control two nodes simultaneously, in other words, there are fewer superiors (x_2) than subordinates (x_1 and x_3). The network in the top right is not controllable because node x_3 is inaccessible for the only input node u_1 in the system. Both systems can be made controllable by adding an extra input node (top).

on a pollinator species, that pollinator tends to depend rather weakly on the plant (and vice versa). We, therefore, create a directed link from species i to species j when $d_{ij} - d_{ji} \geq 0$ to establish the most likely direction of control between a species pair (Figure 1a). Sometimes there is no observed asymmetry between species pairs ($d_{ij} = d_{ji}$), and we cannot infer a dominant direction of control. When this occurs, we deem both species to be equally likely to affect each other and form a reciprocal interaction between them, this is a link from i to j and another from j to i. By basing the direction of the links on the asymmetry of their dependence, we are able to generate a network that is consistent with the dynamics of the community while satisfying the requirements of structural controllability. That allows us to calculate the controllability of the networks and investigate whether there are differences between invaded and uninvaded communities.

Controllability

A system is said to be controllable if it is possible to steer it from an initial to an arbitrary final state within finite time (Kalman, 1963). A simple version of such a system can be described by $\frac{dx}{dt} = Ax + Bu(t)$,

5

where the change of its state over time ($\frac{dx}{dt}$) depends on its current state x (for example, the species' abundances), an external time-varying input *u(t)* (the control signal), and two matrices A and B, which encode information about the network structure and how species respond to external inputs, respectively. In classic control theory, determining wether this system is controllable can be achieved by checking that its controllability matrix $R = [B \quad AB \quad A^2B \quad ... \quad A^{n-1}B]$ has full rank. In complex systems, however, employing this rank condition, or numerical approximations of it, is infeasible because it is hard to fully parameterize A and B (either because the weight of the links changes over time or because they are difficult to measure). Here we use an approach based on the structural controllability theorem (Lin, 1974), which assumes that we are confident about which elements of A and B have either non-zero or zero values (there is an interaction or not), but that we less sure about the precise magnitude of the non-zero values. Using this structural approach we can find out the controllability of a system for every non-zero realisations of the parameters. An intuitive way to understand structural controllability is by looking at its graphical interpretation: from a topological perspective a network is structurally controllable if there are no inaccessible nodes—nodes without incoming links—or dilations—*expansions* of the network (Figure 1b; Supporting Information Section S2).

We are often able to estimate A in ecological networks, as this matrix represents the interactions between species. Part of the control problem resides in estimating a supportable estimation of B, which represents the links between external inputs and species. Naively, any ecological community (and any system for that matter) could be controlled if we control the state of every species independently, but such an approach is typically impractical. Here, we are interested in finding a minimum driver node-set (effectively finding B) with which to make the system controllable. The brute-force search for this minimum driver node-set is computationally prohibitive for most networks as it involves the evaluation of 2^N different controllability matrices. We therefore instead employ a recently-developed approach that shows that the control problem of finding the minimum driver node-set can be mapped into a graph-theoretic problem: maximum matching (Liu & Barabási, 2016; Liu, Slotine, & Barabási, 2011).

Maximum-matching is a widely studied topic in graph theory and is commonly used in multiple applications, ranging from dating apps and wireless communications to organ transplant allocation and peer-to-peer file sharing. A matching in an unweighted directed graph is defined as a set of links that do not share a common start or end nodes; the largest possible matching is called a maximum matching. For example, in a network composed by jobs and job applicants, a matching is any pairing between applicants and positions that satisfies one basic constraint: an applicant can be assigned to at most one position and vice versa. Consequently, a maximum matching is an optimal pairing, one that maximises the number of applicants with jobs and the number of positions filled. Admittedly, the link between matchings and structural controllability may appear far from straightforward. The key is to note that the fundamental
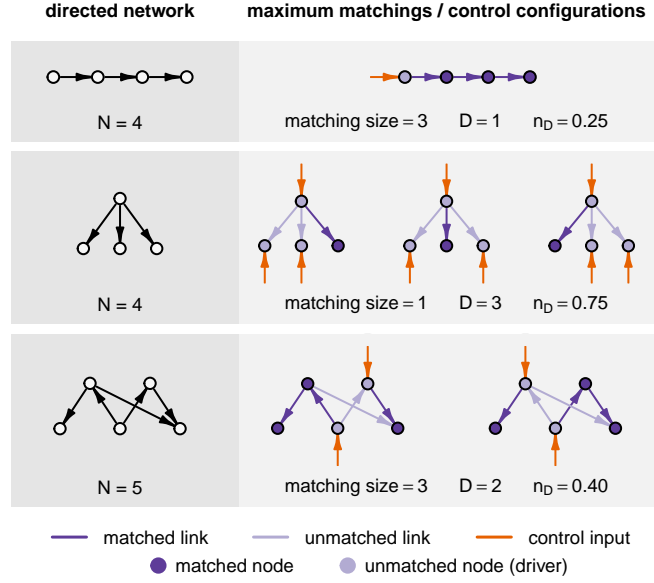
Figure 2: Maximum matchings and control configurations. In directed networks, a maximum matching is the largest possible set of links that do not share start or end nodes (dark purple). Maximum matchings are not necessarily unique, instead, each of them is related to a possible minimum driver node-set in the network (the nodes to which an external control input, in orange, should be applied in order to ensure controllability). The size of the minimum driver node-set D corresponds exactly to the number of unmatched nodes (the number of nodes in the network N minus the matching size). To account for network size, we use the size of the minimum driver node-set relative to the total number of nodes $n_D = D/N$ as a measure of the extent to which the network structure can be harnessed to control the system.

conditions of structural controllability imply that there is a one-to-one relationship between *superior* and *subordinate* nodes just like the one-to-one relationship between jobs and applicants (Figure 1b, bottom left). We thus use the maximum-matching algorithm to find an optimal pairing of superior and subordinate nodes in a manner consistent with the controllability conditions (Supporting Information Section S3.1). Given the result, we can further decompose the matching into a set of paths that reveal how a control signal can flow across the links in a network to reach every node that composes it. As recently shown (Liu et al., 2011), the minimum driver node-set—those to which an external control input should be applied to make the system controllable—corresponds exactly to the *unmatched* nodes in the network (Figure 2).

Differences between invaded and uninvaded networks

Our first ecological objective is to investigate whether the controllability of a community is associated with invasion status or not. Finding out exactly how difficult it is to control a network depends strongly on the particularities of the desired control trajectory as well as the dynamical relationship between nodes. However, we are interested in understanding the controllability of a network in a more general sense, such that it can be applied even when the precise control scenario is known only incompletely. To this end, we chose an indicator that follows directly from our approach: the size of the minimum driver node-set. This

simple metric provides a general indication of how difficult controlling a network might be, as systems that require a large number of external inputs to be fully controlled are intuitively more difficult or costly to manage. For instance, achieving full control in a "network" in which species do not interact at all is relatively difficult as we would require an intervention for every single species. Conversely, the structure of a linear trophic chain can be harnessed to achieve full control using just one intervention targeted the top species; the control signal would then cascade through the trophic levels and reach other species in the community. Specifically, drawing from the structural-controllability literature, we use the size of the minimum driver node-set relative to the total number of species $n_D = \frac{D}{N}$ as a measure of the extent to which the network structure can be harnessed to control the community. In an ecological context, external inputs can be thought of as management interventions that modify the abundance of a particular species.

After finding the minimum driver node-set in each of our networks, we wanted to test whether invasion status or other predictors are correlated to this metric of controllability. We, therefore, use a set of generalised linear models (with binomial error structure). The response variable was the size of the minimum driver node-set n_D of the twenty empirical networks (ten invaded and ten uninvaded). The independent variables evaluated were the invasion status, the network connectance, the network nestedness (NODF) the number of species (since one might naively expect to see a negative relationship between richness and controllability; Menge, 1995), the network asymmetry (an indication of the balance between plant and pollinator diversity), and the interaction strength asymmetry (the asymmetry on the dependences between trophic levels; N. Blüthgen, Menzel, Hovestadt, Fiala, & Blüthgen, 2007). Models were compared using the Akaike Information Criterion for small sample sizes (AICc).

In addition, we also explored whether real networks differ in their architecture from random ones in a concerted way that could impact these results. Specifically, we used two null models each with 99 randomisations per network. In the first, we followed Diego P Vázquez et al. (2007) and maintained the connectance of the network but randomised the visits across species such that the species-specific relative probabilities of interactions were maintained; we then calculated the direction of control and the corresponding size of the minimum driver node-set. For the second null model, we used the empirical directed network and randomly shuffled the direction of control between a species pair.

Species roles

Our second key question is related to how species differ in their ability to drive the population dynamics of the community. Ecologically, these differences are relevant because resources and data are limited, and therefore full control is unfeasible. While calculating the size of the minimum drive node-set can measure the controllability of an ecological community, it does not provide information about the roles that particular species play. To answer our question, we make two kinds of distinctions between species

from a control perspective.

First, we harness the fact there may be multiple maximum matchings for a given network, and each of these maximum matchings indicates unique paths that could potentially be used to control the network. Specifically, we use the frequency with which a species is part of the multiple minimum driver node-sets—its control capacity ϕ—as an estimation of its relative importance in driving the state of the community as a whole (Jia & Barabási, 2013). For example, a species with a control capacity $\phi_i = 1$ is a one that requires external input in every single control configuration. This is, if we were to drive the community to an arbitrary alternate state, we would need to apply a management intervention to this species under *any* optimal control strategy.

To calculate a species control capacity ϕ, we need to enumerate all possible maximum matchings (we detail the algorithm to find all maximum matchings in the Supporting Information Section S3.2). However, enumerating all maximum matchings is extremely expensive from a computational perspective—a network with a couple dozen species could be controlled with several hundred million unique matchings. To solve this problem, we employ a recently-developed algorithm that reveals the control correlations between the nodes in the graph using considerably less computational resources (Zhang, Lv, & Pu, 2016, Supporting Information Section S3.3). Using this algorithm, we are able to identify species that are possible control inputs—those that belong to the minimum driver node-set in at least one of the possible control configurations. One of the original contributions of our work is to extend this algorithm such that it is possible to calculate a highly accurate approximation of the control capacity ϕ_i of every species in the network.

A species control capacity ϕ allow us to identify species that are critical to change or maintain the state of the community. However, in ecological settings, changing the state of the community as a whole is often not required or undesired. Consequently, we make a second distinction between species based on whether they form part of possible control paths in the network and therefore are likely to influence the abundance of another species. We identify these species by checking whether they are classified as *superiors*, this is they are at the beginning of a matched link. While unmatched species correspond to those that should be influenced with an external control input, superior species correspond to those that influence other species internally. Conveniently, unlike the minimum driver node-set, superior species are the same for all control configurations and so in a network without reciprocal links (see below) the probability that a species is superior σ is either 1 or 0.

We calculated the species control capacity ϕ and the probability of being superior σ for each species in our networks. In the networks that contained reciprocal links (because there was no asymmetry in the dependences of a species pair), we averaged a species' ϕ and σ across every possible "non-reciprocal" realisation of the network (more details can be found in the Supporting Information Section S3.4). We

then examined how species-level structural properties were related to these two variables using two sets of generalised mixed-effects models with binomial error structure. As explanatory variables in the fixed component, we included four variables. First, the species contribution to nestedness, which has been proposed as a key feature that promotes stability and robustness in mutualistic networks (S. Saavedra, Stouffer, Uzzi, & Bascompte, 2011). Second, the visitation strength (the sum of a species' interactions), which quantifies the strength of a species associations and is indirectly related to its abundance (Poisot, Canard, Mouquet, & Hochberg, 2012). Third, the direction of asymmetry which quantifies the net balance in dependencies, this is it indicates if a species affects other species more than what they affect it or not (Diego P Vázquez et al., 2007). Fourth, to account for the centrality of a species we included the degree as the fourth fixed effect. To facilitate comparison between variables, all four variables were scaled to have a mean of zero and a standard deviation of one. To identify the models that were best supported by the data we first determined the most parsimonious random structure using the AICc. Specifically, we evaluated a random intercept for observations grouped by the study site, and an intercept and slope for the contribution to nestedness, visitation strength and asymmetry grouped by the trophic guild (plant or pollinator). After identifying the most likely random structure we ranked models from all possible fixed effects combinations using the AICc weight. The relative importance of variables was then assessed by looking at both their effect sizes in the top-ranked models and the cumulative weight of the models in which they are present.

Some species have a control capacity $\phi = 1$. These species are critical to controlling community because they are part of the minimum driver node-set in every control scenario. In other words, it is not possible to change the state of the community to a desired state without managing the abundance of these species. We anticipate that these species have a disproportionally large impact on the community dynamics. Therefore, we determined the critical species in each of the networks and investigated whether they have a larger than average impact on the stable coexistence of species in the community. This property of promoting stable coexistence is called structural stability (R. P. Rohr, Saavedra, & Bascompte, 2014). Mathematically it is denoted by the letter Ω and represents the size of the parameter space (growth rates, carrying capacities) under which all species can sustain positive abundances (S. Saavedra, Rohr, Olesen, & Bascompte, 2016). The contribution of species i to stable coexistence, can be estimated by calculating the structural stability of the community when the focal species i is removed. We then used a t-test to compare the contribution to stable coexistence of critical and non-critical species. More details about the calculation of structural stability can be found in the Supporting Information Section S4.

Finally, we wanted to understand how the control capacity ϕ and the probability of being superior σ is related to metrics of keystoness based on centrality. To do that, for each network, we calculated the Spearman correlation matrix between these two variables and some centrality metrics commonly used

to identify keystone species. Namely degree, betweenness, eigen centrality, page rank, and closeness centrality.

Testing assumptions

Just like the centrality metrics, the information obtained by applying structural controllability depends on the ability of the network to accurately represent the ecological community. We tested the sensitivity of our approach to two fundamental assumptions. First, we tested that visitation is an appropriate proxy to infer interspecific effects. To do that we compared the results obtained using visitation to two alternative metrics in a dataset that lacked invasive species (Ballantyne, Baldock, & Willmer, 2015). Specifically, we also applied the structural controllability approach to networks constructed using pollinator efficiency (which measures the pollen deposition of an interaction) and pollinator importance (which accounts for both pollen deposition and visitation and hence is regarded as a more accurate estimation of the pollination service received by plants; Ne'eman, Jürgens, Newstrom-Lloyd, Potts, & Dafni, 2009). More details in the Supporting Information Section S5. Second, because interspecific dependencies depend on the network topology and consequently on the accurate sampling of interactions, we tested the robustness of structural controllability to the uncertainty involved with the sampling of interactions. To do that we compared the results obtained when using the full network and when randomly removing interactions from the weakest links in the network. This effectively removed the rare interactions from the networks (more details in the Supporting Information Section S6).

Results

The size of the minimum driver node-set relative to the number of species in each network n_D ranged between $n_D = 0.58$ and $n_D = 0.88$ (median 0.74, Figure 3a). We found that the relative size of the minimum driver node-set was not related to any of the independent variables we evaluated. Indeed the null model ($n_D \sim 1$) was the most parsimonious candidate model and all variables were relatively unimportant (based on the AICc; Table S3). The n_D of empirical networks did not differ to that of a null model that roughly preserved the degree distribution and fully preserved the network connectance ($p = 0.48$; Figure 3b). However empirical networks had a larger n_D than null models that preserved the interactions but shuffled the direction of control of the empirical network ($p = 2.4 \times 10^{-7}$).

The models that best explained a species control capacity ϕ, had a random intercept and slope for the species asymmetry grouped by the species' guild (plants or pollinators; Table S4, Table S5). Contribution to nestedness was the most important variable followed by visitation strength (Table 1, Table S6) with
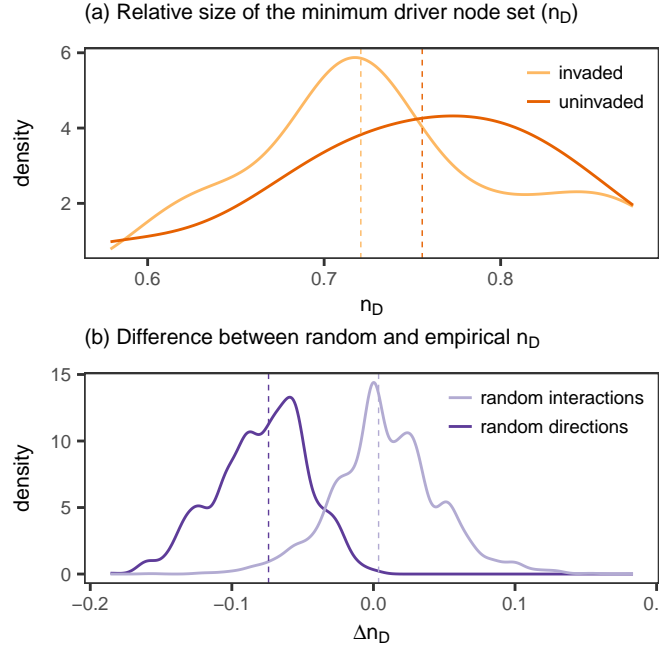
Figure 3: Density plot of (a) the relative size of the minimum driver node-set n_D in the invaded (light) and uninvaded (dark) empirical networks. (b) The density distribution of the difference between the relative size of the minimum driver node-set of random networks and that of empirical networks. We randomised either the species visitation patterns (light line) or randomised the direction of control between a species pair (dark line). The vertical dashed lines indicate the median values of each distribution.

species with large contribution to nestedness and large visitation strength being more likely to have a large control capacity (Figure 4a and b). For the probability of being a superior species σ, the most parsimonuous models did not include any random effect Table S4, Table S5). The most important variables were the species asymmetry and the visitation strength (Table 1, Table S6). Specifically, species with a positive asymmetry (this is they affect their partners more than what their partners affect them), and with a large visitation strength were more likely to be superior (Figure 4c and d). Furthermore, species degree was positively, but only weakly associated with both control capacity and the probability of being a superior species. Remarkably, many species with a low degree, especially pollinators attained a large control capacity in their communities (Figure S10). In addition all invasive plants were both critical ($\phi = 1$) and were superior ($\sigma = 1$) species in their communities.

We also found that critical species (those with a control capacity $\phi = 1$) have a larger impact on species coexistence as the structural stability of the network was considerably reduced when these species were removed from their communities ($p = 3.3 \times 10^{-16}$; Figure 5; Figure S6). Although some pollinators had large values of control capacity, the group of critical species in every community was composed exclusively of plants.

The control capacity of a species ϕ was only weakly correlated with commonly used centrality metrics. The Spearman correlation ranged between -0.14 (with betweeness-centrality) and 0.42 (with eigen-centrality). On the other hand, the probability of being a superior species σ was slightly more similar to established

Table 1: Selection table of the generalised mixed effect models of control capacity ϕ and the probability of being a superior species σ. Only models with a weight larger than 0.01 are shown. All variables were scaled and so the effect sizes are comparable.

	effect sizes						
int.	deg.	asy.	nes.	str.	d.f.	ΔAICc	weight
control capacity models							
0.68	-0.69	—	1.05	1.55	7	0.00	0.35
0.37	—	—	0.95	0.90	6	0.16	0.32
0.79	-0.69	-0.07	1.05	1.55	8	2.03	0.13
0.79	—	-0.24	0.95	0.92	7	2.06	0.12
0.56	—	—	1.11	—	5	4.11	0.04
0.46	0.16	—	1.06	—	6	5.74	0.02
0.61	—	-0.03	1.11	—	6	6.14	0.02
superior probability models							
-5.42	—	7.36	0.38	3.76	4	0.00	0.70
-5.46	-0.13	7.41	0.40	3.78	5	2.35	0.22
-4.92	0.99	6.72	—	3.72	4	4.34	0.08

Terms: intercept (int), degree (deg), asymmetry (asy), visitation strength (str).
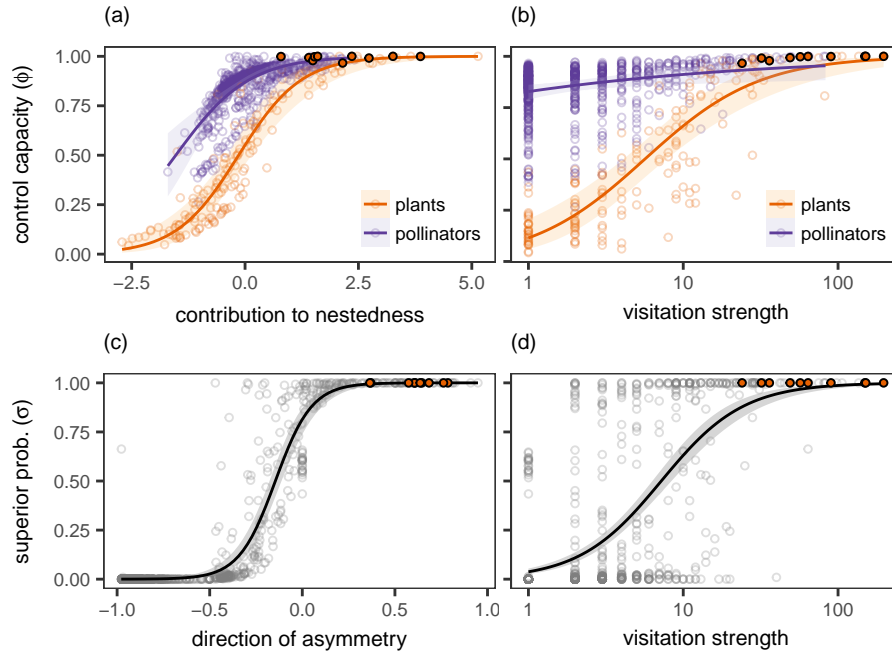

Figure 4: Relationship between the most important variables in the models of control capacity (a, b) and the probability of being a superior species (c, d). The plots show the values predicted by most parsimonious model for each response variable. The model for control capacity include a random effect for the trophic guild , while the model of superior probability does not include any random effect that distinguish between plants and pollinators. The invasive species are depicted with solid orange circles.
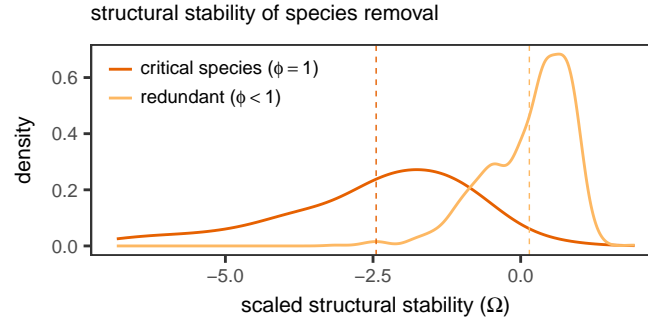
structural stability of species removal

Figure 5: Density plot of the structural stability of the communities after a focal species is removed. The structural stability after critical species have been removed (darker line) is considerable smaller than the average. To allow comparison across communities, the structural stability values were scaled within each network to have a mean of 0 and a standard deviation of one.

312 centrality metrics. The Spearman correlation coefficient ranged between 0.07 (with eigen-centrality) and

313 0.75 (with degree; Figure S11).

314 We found that using visitation as a proxy for the strength of species interactions leads to similar results

315 than those obtained using pollinator importance (regarded as an accurate measure of the pollination

316 service to plants; Supporting Information Section S5; Ne'eman et al., 2009). Importantly, we also found

317 that structural stability is robust to incomplete sampling of interactions. We found a very strong agreement

318 between results obtained using the complete empirical networks and those obtained by randomly removing

319 interactions (Supporting Information Section S6). Despite removing rare interactions and species, the

320 relative size of the minimum driver node-set, the superior species, and the relative rankings of control

321 capacity were generally maintained. Remarkably critical species in the full network were also critical in

322 the vast majority of rarefied networks.

Discussion

References

323 Ballantyne, G., Baldock, K. C. R., & Willmer, P. G. (2015). Constructing more informative plantPollinator

324 networks: Visitation and pollen deposition networks in a heathland plant community. *Proceedings of the*

325 *Royal Society B: Biological Sciences*, *282*(1814), 20151130. doi:10.1098/rspb.2015.1130

326 Bartomeus, I., Vilà, M., & Santamaría, L. (2008). Contrasting effects of invasive plants in plantPollinator

327 networks. *Oecologia*, *155*(4), 761–770. doi:10.1007/s00442-007-0946-1

328 Bascompte, J., & Stouffer, D. B. (2009). The assembly and disassembly of ecological networks. *Philosophical*

Transactions of the Royal Society B: Biological Sciences, *364*(1524), 1781–1787. doi:10.1098/rstb.2008.0226

Bascompte, J., Jordano, P., & Olesen, J. M. (2006). Asymmetric Coevolutionary Networks Facilitate Biodiversity Maintenance. *Science*, *312*(5772), 431–433. doi:10.1126/science.1123412

Blüthgen, N., Menzel, F., Hovestadt, T., Fiala, B., & Blüthgen, N. (2007). Specialization, Constraints, and Conflicting Interests in Mutualistic Networks. *Current Biology*, *17*(4), 341–346. doi:10.1016/j.cub.2006.12.039

Coux, C., Rader, R., Bartomeus, I., & Tylianakis, J. M. (2016). Linking species functional roles to their network roles. *Ecology Letters*, *19*(7), 762–770. doi:10.1111/ele.12612

Dunne, J. A., Williams, R. J., & Martinez, N. D. (2002). Network structure and biodiversity loss in food webs: Robustness increases with connectance. *Ecology Letters*, *5*(4), 558–567. doi:10.1046/j.1461-0248.2002.00354.x

Friedkin, N. E. (1991). Theoretical Foundations for Centrality Measures. *American Journal of Sociology*, *96*(6), 1478–1504.

Guimerà, R., & Amaral, L. A. N. (2005). Cartography of complex networks: Modules and universal roles. *Journal of Statistical Mechanics: Theory and Experiment*, *2005*(02), P02001. doi:10.1088/1742-5468/2005/02/P02001

Isbell, F., & Loreau, M. (2013). Human impacts on minimum subsets of species critical for maintaining ecosystem structure. *Basic and Applied Ecology*, *14*(8), 623–629. doi:10.1016/j.baae.2013.09.001

Jia, T., & Barabási, A.-L. (2013). Control Capacity and A Random Sampling Method in Exploring Controllability of Complex Networks. *Scientific Reports*, *3*(1). doi:10.1038/srep02354

Jordán, F., Benedek, Z., & Podani, J. (2007). Quantifying positional importance in food webs: A comparison of centrality indices. *Ecological Modelling*, *205*(1-2), 270–275. doi:10.1016/j.ecolmodel.2007.02.032

Kaiser-Bunbury, C. N., Muff, S., Memmott, J., Müller, C. B., & Caflisch, A. (2010). The robustness of pollination networks to the loss of species and interactions: A quantitative approach incorporating pollinator behaviour. *Ecology Letters*, *13*(4), 442–452. doi:10.1111/j.1461-0248.2009.01437.x

Kalman, R. E. (1963). Mathematical Description of Linear Dynamical Systems. *Journal of the Society for Industrial and Applied Mathematics Series A Control*, *1*(2), 152–192. doi:10.1137/0301010

Lin, C. T. (1974). Structural Controllability. *IEEE Transactions on Automatic Control*, *19*(3), 201–208. doi:10.1109/TAC.1974.1100557

Liu, Y.-Y., & Barabási, A.-L. (2016). Control principles of complex systems. *Reviews of Modern Physics*,

359 *88*(3). doi:10.1103/RevModPhys.88.035006

360 Liu, Y.-Y., Slotine, J.-J., & Barabási, A.-L. (2011). Controllability of complex networks. *Nature*, *473*(7346),
361 167–173. doi:10.1038/nature10011

362 Lopezaraiza–Mikel, M. E., Hayes, R. B., Whalley, M. R., & Memmott, J. (2007). The impact of an alien
363 plant on a native plantPollinator network: An experimental approach. *Ecology Letters*, *10*(7), 539–550.
364 doi:10.1111/j.1461-0248.2007.01055.x

365 Menge, B. A. (1995). Indirect Effects in Marine Rocky Intertidal Interaction Webs: Patterns and
366 Importance. *Ecological Monographs*, *65*(1), 21–74. doi:10.2307/2937158

367 Mills, L. S., & Doak, D. F. (1993). The Keystone-Species Concept in Ecology and Conservation. *BioScience*,
368 *43*(4), 219–224. doi:10.2307/1312122

369 Motter, A. E. (2015). Networkcontrology. *Chaos*, *25*, 097621. doi:10.1063/1.4931570

370 Ne'eman, G., Jürgens, A., Newstrom-Lloyd, L., Potts, S. G., & Dafni, A. (2009). A framework for com-
371 paring pollinator performance: Effectiveness and efficiency. *Biological Reviews*, no–no. doi:10.1111/j.1469-
372 185X.2009.00108.x

373 Poisot, T., Canard, E., Mouquet, N., & Hochberg, M. E. (2012). A comparative study of ecological
374 specialization estimators: *Species-Level Specialization*. *Methods in Ecology and Evolution*, *3*(3), 537–544.
375 doi:10.1111/j.2041-210X.2011.00174.x

376 Rohr, R. P., Saavedra, S., & Bascompte, J. (2014). On the structural stability of mutualistic systems.
377 *Science*, *345*(6195), 1253497–1253497. doi:10.1126/science.1253497

378 Saavedra, S., Rohr, R. P., Olesen, J. M., & Bascompte, J. (2016). Nested species interactions promote
379 feasibility over stability during the assembly of a pollinator community. *Ecology and Evolution*, *6*(4),
380 997–1007. doi:10.1002/ece3.1930

381 Saavedra, S., Stouffer, D. B., Uzzi, B., & Bascompte, J. (2011). Strong contributors to network persistence
382 are the most vulnerable to extinction. *Nature*, *478*(7368), 233–235. doi:10.1038/nature10433

383 Stouffer, D. B., Sales-Pardo, M., Sirer, M. I., & Bascompte, J. (2012). Evolutionary Conservation of
384 Species' Roles in Food Webs. *Science*, *335*(6075), 1489–1492. doi:10.1126/science.1216556

385 Vázquez, D. P., Melián, C. J., Williams, N. M., Blüthgen, N., Krasnov, B. R., & Poulin, R. (2007). Species
386 Abundance and Asymmetric Interaction Strength in Ecological Networks Author(s): Diego P. Vázquez,
387 Carlos J. Melián, Neal M. Williams, Nico Blüthgen, Boris R. Krasnov and Robert Poulin. *Oikos*, *116*(7),
388 1120–1127. doi:10.1111/j.2007.0030-1299.15828.x

389 Vázquez, D. P., Morris, W. F., & Jordano, P. (2005). Interaction frequency as a surrogate for the total

effect of animal mutualists on plants: Total effect of animal mutualists on plants. *Ecology Letters*, *8*(10), 1088–1094. doi:10.1111/j.1461-0248.2005.00810.x

Woodford, D. J., Richardson, D. M., MacIsaac, H. J., Mandrak, N. E., van Wilgen, B. W., Wilson, J. R. U., & Weyl, O. L. F. (2016). Confronting the wicked problem of managing biological invasions. *NeoBiota*, *31*, 63–86. doi:10.3897/neobiota.31.10038

Zhang, X., Lv, T., & Pu, Y. (2016). Input graph: The hidden geometry in controlling complex networks. *Scientific Reports*, *6*(1). doi:10.1038/srep38209