Journal of Ecology

BRITISH ECOLOGICAL SOCIETY

Keystoneness, centrality, and the structural controllability of ecological networks

SCHOLARONE™
Manuscripts

Keystoneness, centrality, and the 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 B. 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

Abstract

1. An important dimension of a species role is its ability to alter the state and maintain the diversity of its community. Centrality metrics have often been used to identify these species, which are sometimes referred as "keystone" species. However, the relationship between centrality and keystoneness is largely phenomenological and based mostly on our intuition regarding what constitutes an important species. While centrality is useful when predicting which species extinctions could cause the largest change in a community, it says little about how these species could be used to attain or preserve a particular community state.

2. Here we introduce structural controllability, an approach that allows us to quantify the extent to which network topology can be harnessed to achieve a desired state. It also allows us to quantify a species control capacity—its relative importance—and identify the set of species that, collectively, are critical in this context. We illustrate the application of structural controllability with ten pairs of uninvaded and invaded plant-pollinator communities.

3. We found that the controllability of a community is not dependent on its invasion status, but on the asymmetric nature of its mutual dependences. While central species were also likely to have a large control capacity, centrality fails to identify species that, despite being less connected, were critical in their communities. Interestingly, this set of critical species was mostly composed of plants and every invasive species in our dataset was part of it. We also found that species with high control capacity, and in particular critical species, contribute the most to the stable coexistence of their community. This result was true, even when controlling for its degree, abundance/interaction strength, and the relative dependence of their partners.

4. *Synthesis*: Structural controllability is strongly related to the stability of a network and measures the difficulty of managing an ecological community. It also identifies species that are critical to sustain biodiversity and to change or maintain the state of their community and are therefore likely to be very relevant for management and conservation.

Keywords: Keystone species, management interventions, mutualism, network control theory, plant population and community dynamics, species importance, control capacity, structural stability, controllability

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 of 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 and the state of the community—since changes of this nature can have knock-on effects on ecosystem function, diversity, processes, and services (Thompson et al., 2012; Tylianakis, Didham, Bascompte, & Wardle, 2008; Tylianakis, Laliberté, Nielsen, & Bascompte, 2010). This ability is sometimes 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 positional importance that originally stems from social-network research (Friedkin, 1991; Martín González, Dalsgaard, & Olesen, 2010). Generally speaking, central species tend to be better connected and consequently are more likely to participate in the network's "food chains". Because species that participate in more chains are more likely to affect the abundances of other species, centrality metrics have often been used to identify keystone species in the community (Jordán, Benedek, & Podani, 2007). Centrality metrics have been shown to be useful tools to rank species in regard to their potential to alter the abundances of other species, in particular when estimating the probability of secondary extinctions that may follow the loss of a species (Dunne, Williams, & Martinez, 2002; Kaiser-Bunbury, Muff, Memmott, Müller, & Caflisch, 2010).

Despite being conceptually intuitive, the relationship between centrality and a species' presumed impact on the state of the community is largely phenomenological. On the one hand, substantive changes in ecosystem functioning can also occur without complete removal of a species (Mouillot, Graham, Villéger, Mason, & Bellwood, 2013). On the other, we are often interested in a *specific* state of the community that might be desirable to attain (or preserve) because of its biodiversity, resilience, functioning, or the ecosystem services it provides. In these cases, it might be less useful to understand which species may cause *any* change in the community. Instead, we are better served by understanding how the structure of the network can be harnessed to achieve the desired state and which species may play the largest role in this targeted process. When the state of a community is underpinned by more than a single species (often the case in real communities) and we move beyond single-species removals, we might expect the accuracy of centrality to diminish. As a result, community ecology could arguably benefit from an alternative, perhaps more mechanistically-grounded, approach to understand how species affect each other's abundance.

Species abundances—and consequently the state of the community as a whole—are influenced both by the structure of their interactions and the dynamics of these interactions, including the mechanisms of self-regulation (Lever, van Nes, Scheffer, & Bascompte, 2014). However, community and population dynamics can be modelled in innumerable ways, and empirical support for one versus another is often still ambiguous (Holland, DeAngelis, & Bronstein, 2002). The alternative approach should, therefore, ideally acknowledge ecosystem dynamics, but without being overly dependent on the particular choices of how they are characterised. Among the various possibilities *structural controllability*, a branch of 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 (A. E. Motter, 2015). It is inherently designed to deal with system feedbacks and its application has recently been expanded to complex networks (Lin, 1974; Liu & Barabási, 2016). Consistent with long-standing ecological questions, advances in structural controllability 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 (A. E. Motter, 2015).

At its fundamental level, structural controllability first determines whether a system is controllable or not; that is, it asks if a system could ever 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 recently been shown that asking for the controllability of a complex-system is equivalent to finding a particular set of relevant nodes: the set with which is possible to control the state of the whole network (Liu & Barabási, 2016). 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 their *general* ability to modify the system to which they belong.

Here, we apply methods from structural controllability to a particular ecological problem and show how it can be used to generate 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. Moreover, empirical observations indicate that steering the state of some communities—for example during ecosystem restoration or invasive species removal—can be a very difficult task (Woodford et al., 2016). Therefore, we first ask whether there are differences between the controllability of invaded and uninvaded networks. We then expand existing methods from control theory to effectively link the *controllability* of a network with the role of its constituent species. We ask—from a control-theoretic

102 perspective—whether there are key differences between species in the role they play at driving the state

103 of the community and explore the ecological factors related to these differences. This allows us to identify

104 species that might be critical for network control and show that they have a larger than expected impact

105 on the stable coexistence of the community. Finally, we compare the proposed approach to current

106 methods based on species centrality and show how these methods are indeed valuable but ultimately

107 paint a limited picture in regard to the "keystoneness" of a species.

Materials and methods

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

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

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

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

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

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

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

115 30–57 species (mean 38); in contrast to the above, uninvaded communities were obtained from plots

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

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

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

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

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

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

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

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

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

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

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

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

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

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

129 compared to all pollinator visits to plant i. Previous research has shown that mutualistic interactions are
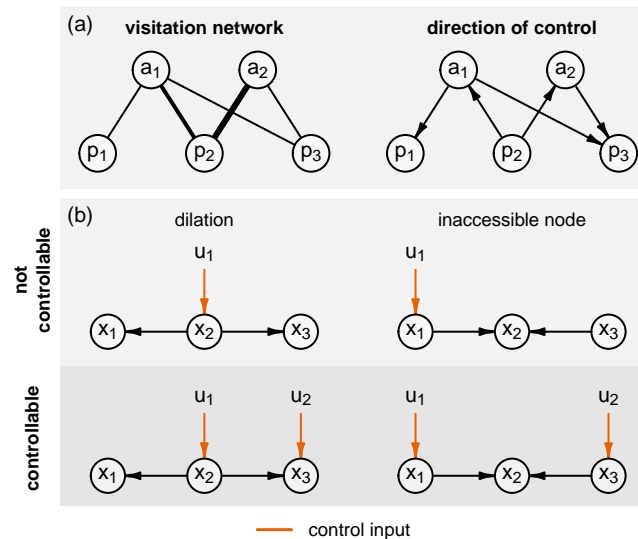
Figure 1: The direction of control and controllability conditions. (a) To establish the direction of control, we start with a weighted visitation network (on the left). In this network, the width of the links corresponds to the frequency of visitation between animals a_i and plants p_i, with wider links indicating more visits. Plant p_1 is visited exclusively by a_1 but p_1 represents only a small fraction of the floral resources exploited by 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 in the control network (on the 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. Any 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 since 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 on 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 (u_2 in both bottom networks).

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

131 on a pollinator species, that pollinator tends to depend rather weakly on the plant (and vice versa).

132 We therefore create a directed link from species i to species j when $d_{ij} - d_{ji} \geq 0$ to establish the most

133 likely direction of control between a species pair (Figure 1a). Sometimes there is no observed asymmetry

134 between species pairs ($d_{ij} = d_{ji}$), and we cannot infer a dominant direction of control. When this occurs,

135 we deem both species to be equally likely to affect each other and leave a reciprocal interaction between

136 them (a link from i to j and another from j to i). By basing the direction of the links on the asymmetry

137 of their dependence, we are able to generate a network that is consistent with the dynamics of the

138 community while satisfying the requirements of structural controllability. This allows us to calculate

139 the controllability of the networks and investigate whether there are differences between invaded and

140 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 $\frac{dx}{dt} = Ax + Bu(t)$, 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 whether 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 parameterise 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 are 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 realisation of the parameters.

We are often able to estimate A in ecological networks, as this matrix represents the interactions between species. Part of the control problem thus 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 where N is the number of species in the community. 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 common start or end nodes; the largest possible matching is called a maximum matching. For example, in a network composed of 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
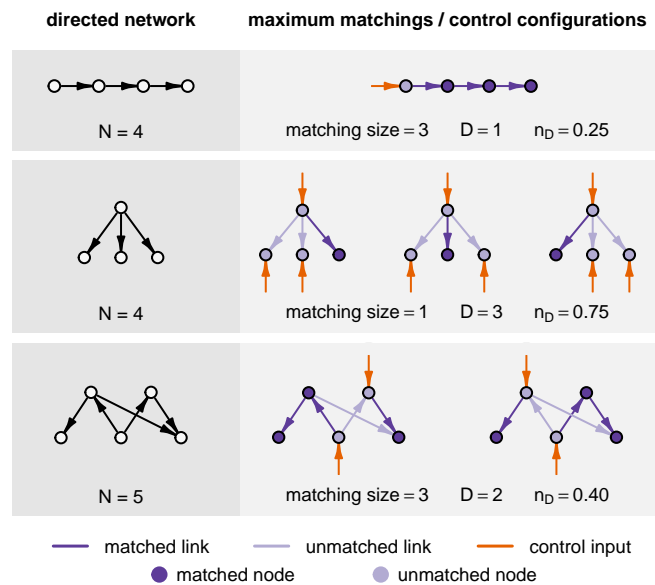
Figure 2: Maximum matchings and control configurations. In directed networks, a maximum matching is the largest possible set of links that do not share start or end nodes (dark purple). Maximum matchings are not necessarily unique; instead, each of them is related to a possible minimum driver-node set 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.

structural controllability may appear far from straightforward.

This link becomes apparent after examining the graphical interpretation of structural controllability: from a topological perspective, a network is structurally controllable if there are no inaccessible nodes—that is, nodes without incoming links—or dilations—*expansions* of the network (Figure 1b; Supporting Information Section S2). The key is to note that these two fundamental 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 (those that can control another node) and subordinate nodes (those that can be controlled by another node) 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 within 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 more 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 at the top species; a suitable control signal could 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 *controllability* of a network—the extent to which the network structure can be harnessed to control the community. The lower n_D the more controllable 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 controllability. We do this using a set of generalised linear models with binomial error structure. The response variable was the relative size of the minimum driver-node set n_D of the twenty empirical networks (ten invaded and ten uninvaded), and we included invasion status as a predictor. As predictors, we also include 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). We compared models 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 relative probabilities of interactions were maintained. We then re-estimated the direction of control and the corresponding

size of the minimum driver-node set, n_D. For the second null model, we used the empirical directed network described above and randomly shuffled the direction of control between a species pair prior to re-estimating the size of the minimum driver-node set.

Species roles

Our second objective is related to how species differ in their ability to drive the population dynamics of the community. We in turn examine whether these differences are also reflected in the role species play at supporting the stable coexistence of other species in the community. Ecologically, these differences are relevant because resources and data are limited, and therefore full control is infeasible. 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 this question, we harness the fact there may be multiple maximum matchings for a given network, and each of these maximum matchings indicates a unique combination of species with which it is possible to control the network. Moreover, some species belong to these combinations more often than do others. We call this property a species "control capacity", ϕ. The higher a species control capacity, the greater the likelihood that they would need to be directly managed to change (or maintain) the ecological state of their community. Therefore, a species control capacity provides an estimation of their relative importance at driving the state of the community (Jia & Barabási, 2013).

To calculate a species control capacity ϕ, we must first enumerate all possible maximum matchings (Supporting Information Section S3.2). Unfortunately, enumerating all maximum matchings is extremely expensive from a computational perspective—a network with a couple dozen species has several hundred million unique maximum matchings. To solve this problem, we employ a recently-developed algorithm that reveals the control correlations between the nodes in the graph while requiring considerably less computational resources (Zhang, Lv, & Pu, 2016). 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. Here, we extend this algorithm such that it is possible to calculate a highly accurate approximation of the control capacity ϕ of every species in the network (Supporting Information Section S3.3). In the networks that contained reciprocal links (because there was no asymmetry in the dependences of a species pair), we averaged a species control capacity ϕ across every possible "non-reciprocal" version of the network (Supporting Information Section S3.4).

We then examined how species-level properties were related to control capacity using a set of generalised linear models with binomial error structure. These models included five predictor variables that mirror the network-level predictors. 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 species strength (the sum of a species' visits), 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; that is, it indicates if a species affects other species more than what they affect it or not (Diego P Vázquez et al., 2007). Fourth, the species degree in order to account for the intrinsic centrality of a species. Finally, we included a categorical variable for the species trophic level (plant or pollinator) and an interaction term between trophic level and the previous four variables. To facilitate comparison between predictors, degree and visitation strength were log-transformed and all four continuous 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. The relative importance of variables was then assessed by looking at their effect sizes in the top-ranked models and the cumulative weight of the models in which they are present.

In addition, we wanted to understand how a species control capacity ϕ described above relates to metrics of keystoneness based on centrality. Specifically, in each network, we calculated the species' degree, betweenness, closeness centrality (Martín González et al., 2010), page rank (McDonald-Madden et al., 2016), and eigen centrality (Jordano, Bascompte, & Olesen, 2006). We then calculated the spearman correlation coefficient between control capacity and each of these centrality metrics.

Our analysis revealed that some species have a control capacity $\phi = 1$. These species are critical to controlling their community because they are part of the minimum driver-node set in *every* control scenario. In other words, it is theoretically impossible to drive the state of the community to a desired state without directly managing the abundance of these species. We thus anticipate that these species have a disproportionally large impact on the community dynamics. To test this hypothesis, we identified these 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. Within mutualistic networks, one useful measure of stable coexistence is called structural stability (R. P. Rohr, Saavedra, & Bascompte, 2014). Mathematically, the structural stability of a network represents the size of the parameter space (i.e., growth rates, carrying capacities, etc.) under which all species can sustain positive abundances (S. Saavedra, Rohr, Olesen, & Bascompte, 2016). The contribution of any given species i to stable coexistence can be estimated by calculating the structural stability of the community when the focal species i is removed. To allow comparison across communities, the structural stability values were scaled within each network to have a mean of zero and a standard deviation of one. Given these species-specific estimates of structural stability, 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

₂₈₆ Section S4.

Testing assumptions

₂₈₇ Just like the centrality metrics, the information obtained by applying structural controllability depends
₂₈₈ on the ability of the network to accurately represent the ecological community. We thus tested the
₂₈₉ sensitivity of our approach to two fundamental assumptions. First, we tested that visitation is an
₂₉₀ appropriate proxy to infer interspecific effects by comparing the results obtained using visitation to two
₂₉₁ alternative metrics in a separate dataset that lacked invasive species (Ballantyne, Baldock, & Willmer,
₂₉₂ 2015). Specifically, we also calculated the controllability (the size of the minimum driver node-set) and the
₂₉₃ control capacity of networks constructed using pollinator efficiency (which measures the pollen deposition
₂₉₄ of an interaction) and pollinator importance (which accounts for both pollen deposition and visitation and
₂₉₅ hence is regarded as a more accurate estimation of the pollination service received by plants; Ne'eman,
₂₉₆ Jürgens, Newstrom-Lloyd, Potts, & Dafni, 2009). More details in the Supporting Information Section S5.

₂₉₇ Second, because interspecific dependencies themselves depend on the network topology and consequently
₂₉₈ on the accurate sampling of interactions, we tested the robustness of structural controllability to the
₂₉₉ uncertainty involved with the sampling of interactions. Here, we compared the results obtained when using
₃₀₀ the full network and when randomly removing interactions from the weakest links in the network. This
₃₀₁ effectively removed the rare interactions from the networks (more details in the Supporting Information
₃₀₂ Section S6).

Results

Controllability

₃₀₃ The size of the minimum driver-node set relative to the number of species in each network n_D ranged
₃₀₄ between $n_D = 0.58$ and $n_D = 0.88$ (median 0.74).

Differences between invaded and uninvaded networks

₃₀₅ We found that the relative size of the minimum driver-node set of invaded communities was not significantly
₃₀₆ different from that of communities that have not been invaded (Figure 3a). In contrast, there was a large
₃₀₇ negative relationship between n_D and the network asymmetry (Figure 3b). Furthermore, there were also
₃₀₈ negative, albeit weaker, relationships between n_D and connectance, nestedness and species richness (Table
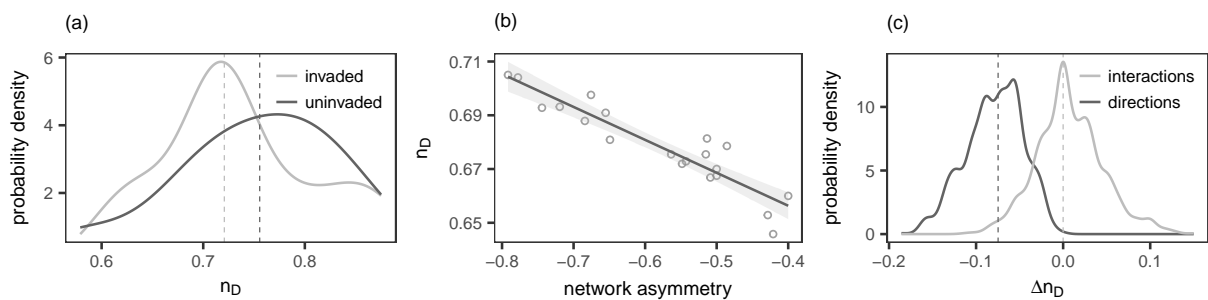
Figure 3: Drivers of network controllability. (a) Probability density of the relative size of the minimum driver-node set n_D in the invaded (light) and uninvaded (dark) empirical networks. (b) Relationship between the asymmetry plant/pollinator richness and n_D. (c) Probability density of the difference 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 in (a) and (c) indicate the median values of the distributions.

309   S3). The relative size of the minimum driver-node set n_D of empirical networks did not differ from that of

310   a null model that roughly preserved the degree distribution and fully preserved the network connectance

311   ($p = 0.66$; Figure 3c). However, empirical networks had a larger n_D than null models that preserved the

312   interactions but shuffled the direction of control of the empirical network ($p = 2.4 \times 10^{-7}$).

Species roles

313   Species varied widely in their control capacity (Figure 4). Pollinators had, in average, larger control

314   capacities than plants. That said, almost no pollinator was critical for network control, where a species

315   is critical for control if it has control capacity $\phi_i = 1$). Plants had a multimodal distribution of control

316   capacity with maxima at both extremes of the distribution (Figure 4a). Intriguingly, every invasive species

317   was critical for network control in each of their communities. The species-level models identified a positive

318   relationship between control capacity ϕ and a species' contribution to nestedness, visitation strength, and

319   the asymmetry of its dependences (Table 1; Figure 5; Table S4). Comparatively, species' degree was only

320   weakly associated with control capacity (Table S5). In fact, many species with a low degree, especially

321   pollinators, exhibited a large control capacity in their communities (Figure S10a).

322   Species control capacity ϕ was only weakly correlated with commonly-used centrality metrics. The

323   Spearman correlation between these ranged between -0.14 (with betweeness-centrality) and 0.42 (with

324   eigen-centrality), see Figure S11a. The correlation coefficient with degree was -0.13, however most species

325   with high degree also tended to attain a high control capacity (Figure S10a).

326   Finally, we found that critical species have a particularly large impact on species coexistence when

327   compared to non-critical species. Indeed, the structural stability of the networks where critical species

328   were removed was considerably lower to those in were non-critical species were removed ($p = 2 \times 10^{-15}$;

329   Figure 6; Supporting Information S4).

Table 1: Selection table of the binomial generalised linear models of species control capacity, ϕ. Only models with a weight larger or equal to 0.01 are shown.

				model terms									
int.	k	l	a	n	s	$k{:}l$	$l{:}a$	$l{:}n$	$l{:}s$	d.f.	ΔAICc	weight
-1.20		+	0.80	0.15	0.29		+	+		7	0.00	0.48
-1.19		+	0.76	0.13	0.35		+	+	+	8	1.52	0.22
-1.26	-1.24	+	1.44	0.39	1.07	+	+		+	9	4.09	0.06
-1.37	-0.66	+	1.03		1.06	+	+		+	8	4.39	0.05
-1.27	-1.15	+	1.37	0.33	1.07	+	+	+	+	10	4.92	0.04
-1.37	-0.10	+	0.90		0.43	+	+			7	6.36	0.02
-1.25	-0.28	+	1.24	0.40		+	+			7	6.47	0.02
-1.24	-0.62	+	1.29	0.38	0.40	+	+			8	6.50	0.02
-1.39	0.30	+	0.83			+	+			6	6.72	0.02
-1.28	-0.17	+	1.16	0.32		+	+	+		8	7.03	0.01
-1.26	-0.53	+	1.23	0.32	0.39	+	+	+		9	7.42	0.01
-1.02		+	0.69	0.30	0.31		+			6	7.48	0.01

Terms: intercept (int), degree (k), trophic level (l), asymmetry (a), contribution to nestedness (n), visitation strength (s).
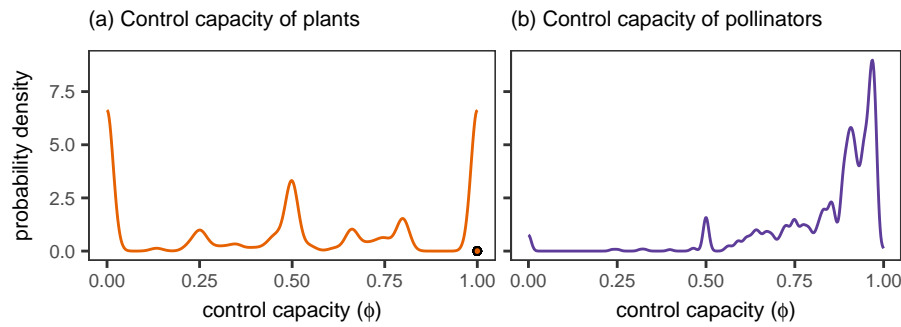

Figure 4: Probability density of the control capacity ϕ of (a) plants and (b) pollinators across all networks. The control capacity of all invasive species is $\phi = 1$ and is depicted with solid circles.
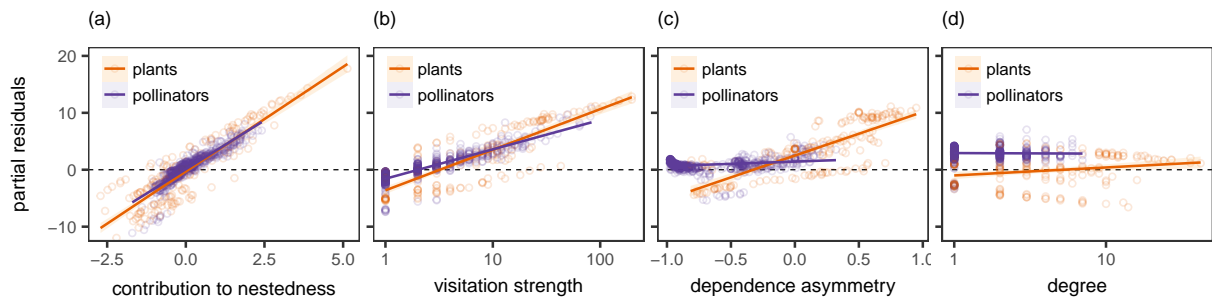

Figure 5: Partial-residual plots for the independent variables: (a) contribution to nestedness, (b) visitation strength, (c) asymmetry of dependences, and (d) degree. Partial-residual plots show the relationship between control capacity and each of the independent variables while acccounting for all other remaining variables. Ploted values correpond to the predictions of the weighted average of the models shown in Table 1.
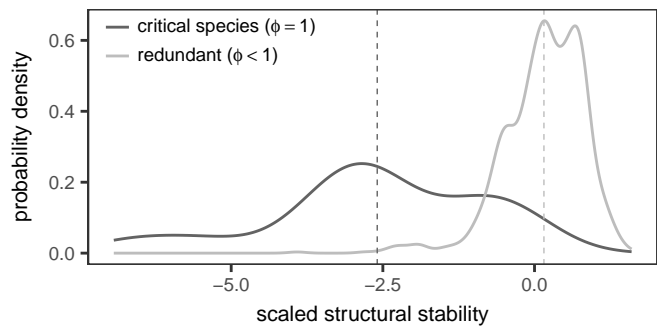
Figure 6: Probability density of the structural stability of the communities after a single focal species is removed. Mathematically, the structural stability of a network represents the size of the parameter space (i.e., growth rates, carrying capacities, etc.) under which all species can sustain positive abundances. The structural stability of communities in which critical species have been removed (darker line) is considerably smaller than that of communities in which non-critical species have been removed. This indicates that critical species contribute more to the stable coexistence of their communities. To allow comparison across communities, the structural stability values were scaled within each network to have a mean of zero and a standard deviation of one. Here, we assume values of the mutualistic trade-off and mean interspecific competition of $\delta = 0$ and $\rho = 0.01$ respectively (S. Saavedra et al., 2016). However, the choice of these parameters does not affect the results (Supporting Information S4).

Testing assumptions

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

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

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

333 structural stability is robust to incomplete sampling of interactions. Indeed, we found strong agreement

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

335 the weakest interactions (Supporting Information Section S6). Despite removing rare interactions and

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

337 of control capacity were generally maintained. Of particular note, we found that critical species in the full

338 network were also critical in the vast majority of rarefied networks.

Discussion

339 Our main goal was to understand the role that species play at both modifying the abundance of the

340 species they interact with and the state of the community as a whole. To achieve that goal we applied

341 *structural controllability*, a field at the intersection between control and complex theory that allow us

342 to obtain two key pieces of information: the *controllability* of a network and a species *control capacity*

343 (Table 2). We found that the controllability of a network does not depend on its invasion status and that

344 the species that are critical to altering the state of the community are also the ones that most sustain the

345 stable coexistence of species in their communities.

Table 2: Glossary

network control
    A network is said to be controllable if it is possible to steer it from an initial to an arbitrary final state within finite time.

controllability
    The intrinsic difficulty of controlling an ecological community. It is measured by the relative size of the minimum driver-node set, n_D. It also indicates the extent to which network structure can be harnessed for network control.

minimum driver-node set
    One of the sets of species whose abundances need to be directly managed in order to achieve full control of the community. The minimum driver-node sets can be obtained by finding all maximum matchings in a network.

maximum matching
    A matching is a set of links that do not share any common start or end nodes; the largest possible matching is called a maximum matching.

control configuration
    One of the species combinations with which is possible to achieve network control. Optimal control configurations are given by the minimum driver-node sets.

control capacity
    The relative frequency ϕ which with a species is part of the optimal control configurations of a network.

critical species
    A species with a maximal control capacity $\phi = 1$

superior node
    A species is a superior node if it can internally affect the abundance of other species in the network. Superior nodes make up the chains that propagate the control signals through the network.

346 Our results indicate that fully controlling ecological networks might currently be out of reach for all

347 but the smallest communities (A. E. Motter, 2015). Indeed, the median size of the relative minimum

348 driver-node set in our dataset was $n_D = 0.74$, a high value when compared to other complex systems

349 in which controllability has been investigated (the lower n_D the more controllable the community). For

350 instance, only gene regulation networks appear to achieve similar levels of controllability while most

351 social, power transmission, Internet, neuronal, and even metabolic networks seem to be "easier" to control

352 $(0.1 < n_D < 0.35)$ (Liu et al., 2011). Structural controllability provides solid theoretical rationale for the

353 many difficulties encountered in the management and restoration of natural communities (Woodford et

354 al., 2016). Nevertheless, structural controllability might be helpful at identifying communities in which

355 changes in the ecological state are more likely to occur

356 The differences between the controllability across networks are likely to arise from differences in their

357 structure rather than their invasion status. Specifically, when controlling for network structure, we found

358 no difference between the controllability of invaded and uninvaded networks. instead controllability is

359 almost completely constrained by the patterns of species richness at each trophic guild and their degree

360 distributions (N. Blüthgen et al., 2007; C. J. Melián & Bascompte, 2002). These two factors are particularly

361 relevant because they govern the asymmetric nature of mutual dependences, which themselves provide

362 the foundation of structure and stability in mutualistic networks (Astegiano, Massol, Vidal, Cheptou, &

363 Guimarães, 2015; J. Bascompte et al., 2006; J. Memmott, Waser, & Price, 2004).

364 Accordingly, our results suggest that structural controllability is closely related to the dynamic persistence

365 of an ecological community based on two lines of evidence. First, we found a comparatively small

366 but thought-provoking negative relationship between the controllability of a network and its nestedness.

367 Previous studies indicate that nestedness promotes species coexistence and confers robustness to extinction

368 (Bastolla et al., 2009; J. Memmott et al., 2004) even at the expense of the dynamic stability of the

369 mutualistic community (S. Saavedra et al., 2016). These observations are in agreement with our results,

370 as we would expect the dynamic stability of a community to be correlated to the difficulty to control it.

371 Second, species' control capacity was strongly correlated to their contribution to nestedness and critical

372 species had the largest impact to the stable coexistence of species in their communities. Therefore, species

373 that play a key role at determining the state of the community might also be more key to "maintain

374 the organization and diversity of their ecological communities", one of the hallmarks of keystone species

375 (Mills & Doak, 1993).

376 When controlling for a species' strength, which is indirectly a proxy of its abundance, and the net balance

377 of its dependencies, we found that control capacity could not be easily predicted by species' degree or

378 other metrics of centrality. For instance, some species with a low degree achieved the maximum control

379 capacity and were critical for control in their communities. At first glance, our findings challenge numerous

studies that highlight the role that central species play in the dynamics of their communities and their utility at predicting species extinctions (Jordan, 2009). However, further, inspection shows that our results do not contradict these findings; most species with a large degree also have a large control capacity and all of them were classified as superior nodes which corroborates the utility of classic centrality metrics. Putting these observations together, our results therefore take previous findings one step further and suggest that centrality might paint an incomplete picture of the relevance of species.

Other conceptual differences between structural controllability and centrality metrics provide three key insights into the conservation of ecological networks. First, structural controllability emphasizes that the effect a species has on the abundance of other species is not independent of the effects of other other species in their community. The rankings provided by centrality metrics and other heuristics fail to account for the collective influence of several species at once. Second, it demonstrates that to ensure the persistence of a community it is often necessary to consider the abundances of more than a single species, even when full control is infeasible or undesired (for example 90% of our communities contained more than one critical species). Third, structural controllability explicitly acknowledges the existence of multiple management strategies and some will be better than others depending on the context. Approaches to prioritise species for conservation and reintroduction based on traits or centrality are still useful and are likely to overlap with species control capacity (Devoto, Bailey, Craze, & Memmott, 2012; Pires, Marquitti, & Guimarães, 2017). Stepping back, our results also provide support to the idea that management decisions should not be based on a single technique but indicate that focusing on ecosystem processes and interactions may be more effective than traditional ranking-based approaches (Harvey, Gounand, Ward, & Altermatt, 2017).

Our choice of studying invaded/uninvaded networks was based on a desire to contrast the extensive empirical evidence of the role of invasive plants with our theoretical results. We found that invasive plants were always critical for network control and as such our results were in line with our expectations. Invasive plants have been previously found to exacerbate the asymmetries in their communities (Aizen, Morales, & Morales, 2008; Bartomeus et al., 2008; Henriksson, Wardle, Trygg, Diehl, & Englund, 2016) and to be central in their communities (Palacio, Valderrama-Ardila, & Kattan, 2016; Vila et al., 2009). We found that invasive plants are, however, unlikely to be inherently different from their native counterparts (Emer, Memmott, Vaughan, Montoya, & Tylianakis, 2016; Stouffer, Cirtwill, & Bascompte, 2014). Just like any other mutualist in our dataset, invasive plants tended to attain a high control capacity proportional to the degree to which they contribute to network persistence, are abundant, and depend little on other species. Furthermore, our observation that changes in the abundance of invasive plants (and presumably all critical species) are crucial to modify the state of the community agrees with recent evidence showing that ecosystem restoration focused on the eradication of invasive plants can have transformative desirable effects in plant-pollinator communities (Kaiser-Bunbury et al., 2017). However, our results also suggest

that removals must be exercised with caution. Not only it is hard to predict the direction in which the system will change, but we also show that critical species can underpin the coexistence of species and therefore some communities may be acutely vulnerable to their eradication (Albrecht, Padron, Bartomeus, & Traveset, 2014; Traveset et al., 2013).

Structural controllability assumes that the networks can be approximated using linear functional responses (Liu & Barabási, 2016). The ramifications of this assumption imply that, while structural controllability is useful to identify species that are relevant for network control, it cannot be used to design the *exact* interventions that should be applied to these species in order to achieve a desired state. In an ideal scenario, we would completely incorporate the species dynamics into the controllability analysis (Cornelius, Kath, & Motter, 2013); the reality is that such information is rarely available in most ecological scenarios. In contrast, structural controllability only requires a quantitative approximation of the network's interactions to gain valuable insight from the community. Furthermore, while the relationship between centrality and keystoneness is based on an intuitive understanding of what a keystone species is, the assumptions of structural controllability are explicit and the estimation of a species importance arises from a mechanistic understanding of the population dynamics between species. By accounting for network dynamics (even if in a simple way), structural stability incorporates more ecological realism, especially in the extreme scenario in which the state of a community is only marginally affected by the topology of their interactions.

Conclusions

Here we show that structural controllability can be applied in an ecological setting to gain insight into the stability of a community and the role that species play at modifying the abundance of other species and ultimately the state of the community. These characteristics make structural stability an ideal framework to evaluate the effects of invasions and other types of perturbations. Importantly, structural controllability can be used to identify critical species in the community that promote biodiversity and underpin the stable coexistence of species in their community. Collectively, critical species dominate the state of their community and therefore are likely to be highly relevant for ecosystem management and conservation. While useful, centrality metrics, which have often been used as a proxy for keystoneness, fail to identify some of these species, highlighting their limitations when we fully embrace the notion that ecological communities are dynamical systems. Paine (1969) showed nearly 50 years ago that one single species can dominate the state of its community. Structural controllability suggests that this situation might be the exception rather than the rule. We see our study as a starting point to study the controllability of ecological and socio-ecological systems where many exciting questions lie ahead.

Acknowledgements

Author contributions

453   DBS conceived the idea; all authors contributed to the development of the theoretical framework. EFC

454   performed all analysis. EFC and DBS wrote the manuscript. All authors contributed to its revision.

Data accessibility

455   All data used in this manuscript have already been published by Lopezaraiza–Mikel et al. (2007),

456   Bartomeus et al. (2008), and Ballantyne et al. (2015) The reader should refer to the original sources to

457   access the data.

References

458   Aizen, M. A., Morales, C. L., & Morales, J. M. (2008). Invasive Mutualists Erode Native Pollination

459   Webs. *PLoS Biology*, *6*(2), e31. doi:10.1371/journal.pbio.0060031

460   Albrecht, M., Padron, B., Bartomeus, I., & Traveset, A. (2014). Consequences of plant invasions on

461   compartmentalization and species' roles in plant-pollinator networks. *Proceedings of the Royal Society B:*

462   *Biological Sciences*, *281*(1788), 20140773–20140773. doi:10.1098/rspb.2014.0773

463   Astegiano, J., Massol, F., Vidal, M. M., Cheptou, P.-O., & Guimarães, P. R. (2015). The Robustness

of Plant-Pollinator Assemblages: Linking Plant Interaction Patterns and Sensitivity to Pollinator Loss. *PLOS ONE*, *10*(2), e0117243. doi:10.1371/journal.pone.0117243

Ballantyne, G., Baldock, K. C. R., & Willmer, P. G. (2015). Constructing more informative plantPollinator networks: Visitation and pollen deposition networks in a heathland plant community. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1814), 20151130. doi:10.1098/rspb.2015.1130

Bartomeus, I., Vilà, M., & Santamaría, L. (2008). Contrasting effects of invasive plants in plantPollinator networks. *Oecologia*, *155*(4), 761–770. doi:10.1007/s00442-007-0946-1

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

Bastolla, U., Fortuna, M. A., Pascual-García, A., Ferrera, A., Luque, B., & Bascompte, J. (2009). The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature*, *458*(7241), 1018–1020. doi:10.1038/nature07950

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

Cornelius, S. P., Kath, W. L., & Motter, A. E. (2013). Realistic control of network dynamics. *Nature Communications*, *4*, 1942. doi:10.1038/ncomms2939

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

Devoto, M., Bailey, S., Craze, P., & Memmott, J. (2012). Understanding and planning ecological restoration of plant-pollinator networks: Understanding network restoration. *Ecology Letters*, *15*(4), 319–328. doi:10.1111/j.1461-0248.2012.01740.x

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

Emer, C., Memmott, J., Vaughan, I. P., Montoya, D., & Tylianakis, J. M. (2016). Species roles in plant-pollinator communities are conserved across native and alien ranges. *Diversity and Distributions*, *22*(8), 841–852. doi:10.1111/ddi.12458

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

Harvey, E., Gounand, I., Ward, C. L., & Altermatt, F. (2017). Bridging ecology and conservation: From ecological networks to ecosystem function. *Journal of Applied Ecology*, *54*(2), 371–379. doi:10.1111/1365-2664.12769

Henriksson, A., Wardle, D. A., Trygg, J., Diehl, S., & Englund, G. (2016). Strong invaders are strong defenders - implications for the resistance of invaded communities. *Ecology Letters*, *19*(4), 487–494. doi:10.1111/ele.12586

Holland, J. N., DeAngelis, D. L., & Bronstein, J. L. (2002). Population Dynamics and Mutualism: Functional Responses of Benefits and Costs. *The American Naturalist*, *159*(3), 231–244. doi:10.1086/338510

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

Jordan, F. (2009). Keystone species and food webs. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1524), 1733–1741. doi:10.1098/rstb.2008.0335

Jordano, P., Bascompte, J., & Olesen, J. M. (2006). The ecological consequences of complex topology and nested structure in pollination webs. In N. M. Waser & J. Ollerton (Eds.), *Plant-Pollinator Interactions: From Specialization to Generalization* (pp. 173–199). University of Chicago Press.

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., Mougal, J., Whittington, A. E., Valentin, T., Gabriel, R., Olesen, J. M., & Blüthgen, N. (2017). Ecosystem restoration strengthens pollination network resilience and function. *Nature*, *542*(7640), 223–227. doi:10.1038/nature21071

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*

525 *Industrial and Applied Mathematics Series A Control*, *1*(2), 152–192. doi:10.1137/0301010

526 Lever, J. J., van Nes, E. H., Scheffer, M., & Bascompte, J. (2014). The sudden collapse of pollinator
527 communities. *Ecology Letters*, *17*(3), 350–359. doi:10.1111/ele.12236

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

530 Liu, Y.-Y., & Barabási, A.-L. (2016). Control principles of complex systems. *Reviews of Modern Physics*,
531 *88*(3). doi:10.1103/RevModPhys.88.035006

532 Liu, Y.-Y., Slotine, J.-J., & Barabási, A.-L. (2011). Controllability of complex networks. *Nature*, *473*(7346),
533 167–173. doi:10.1038/nature10011

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

537 Martín González, A. M., Dalsgaard, B., & Olesen, J. M. (2010). Centrality measures and the
538 importance of generalist species in pollination networks. *Ecological Complexity*, *7*(1), 36–43.
539 doi:10.1016/j.ecocom.2009.03.008

540 McDonald-Madden, E., Sabbadin, R., Game, E. T., Baxter, P. W. J., Chadès, I., & Possingham,
541 H. P. (2016). Using food-web theory to conserve ecosystems. *Nature Communications*, *7*, 10245.
542 doi:10.1038/ncomms10245

543 Melián, C. J., & Bascompte, J. (2002). Complex networks: Two ways to be robust?: Complex networks:
544 Two ways to be robust? *Ecology Letters*, *5*(6), 705–708. doi:10.1046/j.1461-0248.2002.00386.x

545 Memmott, J., Waser, N. M., & Price, M. V. (2004). Tolerance of pollination networks to species extinctions.
546 *Proceedings of the Royal Society B: Biological Sciences*, *271*(1557), 2605–2611. doi:10.1098/rspb.2004.2909

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

549 Mills, L. S., & Doak, D. F. (1993). The Keystone-Species Concept in Ecology and Conservation. *BioScience*,
550 *43*(4), 219–224. doi:10.2307/1312122

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

552 Mouillot, D., Graham, N. A., Villéger, S., Mason, N. W., & Bellwood, D. R. (2013). A functional
553 approach reveals community responses to disturbances. *Trends in Ecology & Evolution*, *28*(3), 167–177.
554 doi:10.1016/j.tree.2012.10.004

555 Ne'eman, G., Jürgens, A., Newstrom-Lloyd, L., Potts, S. G., & Dafni, A. (2009). A framework for com-

556 paring pollinator performance: Effectiveness and efficiency. *Biological Reviews*, no–no. doi:10.1111/j.1469-
557 185X.2009.00108.x

558 Paine, R. T. (1969). A Note on Trophic Complexity and Community Stability. *The American Naturalist*,
559 *103*(929), 91–93.

560 Palacio, R. D., Valderrama-Ardila, C., & Kattan, G. H. (2016). Generalist Species Have a Central Role
561 In a Highly Diverse Plant-Frugivore Network. *Biotropica*, *48*(3), 349–355. doi:10.1111/btp.12290

562 Pires, M. M., Marquitti, F. M., & Guimarães, P. R. (2017). The friendship paradox in species-rich
563 ecological networks: Implications for conservation and monitoring. *Biological Conservation*, *209*, 245–252.
564 doi:10.1016/j.biocon.2017.02.026

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

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

570 Saavedra, S., Rohr, R. P., Olesen, J. M., & Bascompte, J. (2016). Nested species interactions promote
571 feasibility over stability during the assembly of a pollinator community. *Ecology and Evolution*, *6*(4),
572 997–1007. doi:10.1002/ece3.1930

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

575 Stouffer, D. B., Cirtwill, A. R., & Bascompte, J. (2014). How exotic plants integrate into pollination
576 networks. *Journal of Ecology*, *102*(6), 1442–1450. doi:10.1111/1365-2745.12310

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

579 Thompson, R. M., Brose, U., Dunne, J. A., Hall, R. O., Hladyz, S., Kitching, R. L., . . . Tylianakis, J. M.
580 (2012). Food webs: Reconciling the structure and function of biodiversity. *Trends in Ecology & Evolution*,
581 *27*(12), 689–697. doi:10.1016/j.tree.2012.08.005

582 Traveset, A., Heleno, R., Chamorro, S., Vargas, P., McMullen, C. K., Castro-Urgal, R., . . . Olesen,
583 J. M. (2013). Invaders of pollination networks in the Galapagos Islands: Emergence of novel com-
584 munities. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1758), 20123040–20123040.
585 doi:10.1098/rspb.2012.3040

586 Tylianakis, J. M., Didham, R. K., Bascompte, J., & Wardle, D. A. (2008). Global change and species inter-

587 actions in terrestrial ecosystems. *Ecology Letters*, *11*(12), 1351–1363. doi:10.1111/j.1461-0248.2008.01250.x

588 Tylianakis, J. M., Laliberté, E., Nielsen, A., & Bascompte, J. (2010). Conservation of species interaction

589 networks. *Biological Conservation*, *143*(10), 2270–2279. doi:10.1016/j.biocon.2009.12.004

590 Vázquez, D. P., Melián, C. J., Williams, N. M., Blüthgen, N., Krasnov, B. R., & Poulin, R. (2007). Species

591 Abundance and Asymmetric Interaction Strength in Ecological Networks Author(s): Diego P. Vázquez,

592 Carlos J. Melián, Neal M. Williams, Nico Blüthgen, Boris R. Krasnov and Robert Poulin. *Oikos*, *116*(7),

593 1120–1127. doi:10.1111/j.2007.0030-1299.15828.x

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

595 effect of animal mutualists on plants: Total effect of animal mutualists on plants. *Ecology Letters*, *8*(10),

596 1088–1094. doi:10.1111/j.1461-0248.2005.00810.x

597 Vila, M., Bartomeus, I., Dietzsch, A. C., Petanidou, T., Steffan-Dewenter, I., Stout, J. C., & Tscheulin, T.

598 (2009). Invasive plant integration into native plant-pollinator networks across Europe. *Proceedings of the*

599 *Royal Society B: Biological Sciences*, *276*(1674), 3887–3893. doi:10.1098/rspb.2009.1076

600 Woodford, D. J., Richardson, D. M., MacIsaac, H. J., Mandrak, N. E., van Wilgen, B. W., Wilson, J. R.

601 U., & Weyl, O. L. F. (2016). Confronting the wicked problem of managing biological invasions. *NeoBiota*,

602 *31*, 63–86. doi:10.3897/neobiota.31.10038

603 Zhang, X., Lv, T., & Pu, Y. (2016). Input graph: The hidden geometry in controlling complex networks.

604 *Scientific Reports*, *6*(1). doi:10.1038/srep38209

Keystoneness, centrality, and the structural controllability of ecological networks

Supporting Information

Fernando Cagua, Kate L. Wootton, Daniel B. Stouffer

S1   Characteristics of the empirical networks

The networks studied had species richness ranging between 19 and 87 when considering only the largest component in each network). As shown by the network asymmetry AS (N. Blüthgen, Menzel, Hovestadt, Fiala, & Blüthgen, 2007), the networks had a low ratio of plants to pollinators overall. Furthermore, the networks had relatively low levels of nestedness (when measured using the quantitative version of the NODF index; Almeida-Neto & Ulrich, 2011). Details for each network can be found in the Table S1.

Table S1: Properties of the analysed plant-pollinator communities. Here show the number of species (n_s), the number of plants (n_p), the number of pollinators (n_a), the network connectance (c), the network asymmetry (AS), and the network nestedness (NODF index). All properties correspond to the network's largest component. British networks were assembled by Lopezaraiza-Mikel et al. (2007), Spanish were networks assembled by Bartomeus et al. (2008).

site	invader	n_s	n_p	n_a	c	AS	NODF	location
1	—	35	9	26	0.17	-0.49	8.68	Cap de Creus, Spain
1	*Carpobrotus affine*	57	10	47	0.17	-0.65	13.27	Cap de Creus, Spain
2	—	40	10	30	0.16	-0.50	11.66	Cap de Creus, Spain
2	*Carpobrotus affine*	38	11	27	0.21	-0.42	15.04	Cap de Creus, Spain
3	—	31	7	24	0.19	-0.55	12.91	Cap de Creus, Spain
3	*Opuntia stricta*	33	8	25	0.18	-0.52	9.96	Cap de Creus, Spain
4	—	35	10	25	0.17	-0.43	12.43	Cap de Creus, Spain
4	*Carpobrotus affine*	57	14	43	0.14	-0.51	13.70	Cap de Creus, Spain
5	—	35	8	27	0.19	-0.54	11.91	Cap de Creus, Spain
5	*Opuntia stricta*	32	8	24	0.19	-0.50	10.96	Cap de Creus, Spain
6	—	30	9	21	0.17	-0.40	6.91	Cap de Creus, Spain
6	*Opuntia stricta*	37	9	28	0.17	-0.51	12.45	Cap de Creus, Spain
7	—	37	6	31	0.19	-0.68	18.33	Bristol, United Kingdom
7	*Impatients grandulifera*	57	8	49	0.20	-0.72	14.36	Bristol, United Kingdom
8	—	48	5	43	0.21	-0.79	5.87	Bristol, United Kingdom
8	*Impatients grandulifera*	87	15	72	0.11	-0.66	8.12	Bristol, United Kingdom
9	—	55	12	43	0.13	-0.56	13.01	Bristol, United Kingdom
9	*Impatients grandulifera*	86	11	75	0.13	-0.74	13.40	Bristol, United Kingdom
10	—	19	3	16	0.38	-0.68	9.99	Bristol, United Kingdom
10	*Impatients grandulifera*	54	6	48	0.21	-0.78	7.56	Bristol, United Kingdom

S2   Structural controllability

Structural controllability avoids the limitation of not knowing the exact values of the matrices A and B. Structural controllability boils down to two conditions: a system is controllable if there are no inaccessible nodes or dilations. A node is inaccessible if there are no directed paths between it and the input nodes. "Dilations are subgraphs in which a small subset of nodes attempts to rule a larger subset of nodes. In other words, there are more 'subordinates' than 'superiors' " (Liu & Barabási, 2016).

The goal of structural controllability is to use the information contained in **A** to generate a supportable estimate of **B**. This focus allows us to gain insight of the inherent controllability of a network, and the roles of the species that compose it, without being overly dependent on the particular choices of how the system dynamics are modelled or characterised. The trade-off of this approach is that, because of the assumption of linearity, structural controllability alone does not allow us to fully design the time-varying control signal $u(t)$ that can drive the system from one particular equilibrium to another. Nevertheless, the lessons gained when assuming linearity—at both the network and the species level—are a prerequisite for eventually understanding nonlinear control (Liu & Barabási, 2016; Liu, Slotine, & Barabási, 2011).

S3   Maximum matching

S3.1 Finding a single maximum matching

Our approach to finding the minimum number of driver nodes relies on finding maximum matchings. We start with a directed network in which the direction of the link represents the direction of control (Figure S1 left panel). We then construct an alternative representation of the directed network in which each node of the directed network is represented by two nodes that indicate their outgoing and incoming links respectively (Figure S1 centre panel). Finding a maximum matching in this alternative representation is equivalent to finding the largest possible set of edges in which one node on the left-hand side is connected to at most one node on the right-hand side. To find the maximum matching we use the push-relabel algorithm implemented in `max_bipartite_matching` in the R package igraph 1.0.1 (Csardi & Nepusz, 2006). Once we have the matching (shown in the Figure S1 center panel) it is then easy to identify the roles of each node in this representation: nodes on the top-level that are connected to a matched link (dark purple) are superior while those connected to a matched link on the bottom-level are matched. This information can then be mapped back to the original representation to identify the control paths and the driver nodes in the network (Figure S1 right panel).

To further illustrate our methodology here, we also show the approach for the smallest of our empirical networks, the uninvaded network at site 10 (Table S1; Figure S2). This network is composed of 19 species of which three are non-invasive plants and the other 16 are pollinators. The one-to-one relationship between matched and superior nodes implies that in order to achieve full network controllability, most pollinators would be unmatched, and hence are classified as driver nodes that require external intervention. At the same time, both plants in the community, *Heracleum sphondylum* and *Rubus fructicosus*, and one of the pollinators, *Orthotylus/Lygocorus*, tend to be classified as superior nodes.

S3.2 Finding all possible maximum matchings

The algorithm implemented in `max_bipartite_matching`, however, is only able to find **one** of possibly many maximum matchings in a network. Though one maximum matching is enough to calculate n_D and hence to provide an indication of the manageability of a community, it is not sufficient to estimate the role of individual species. To do that, we need to calculate all possible maximum matchings. To do this, we again start from the alternative bipartite representation in Figure S1b and assign an identity to each of the links in the network. We will call this bipartite representation *P*. We then construct the line graph of the alternative bipartite representation *L(P)* (Figure S3). Each node in *L(P)* represents a link in *P* and these are connected to each other if and only if they share a common node in *P*. We then calculate *H*, the complement graph of *L(P)* and identify all of its maximal cliques (Figure S3). Here some extra definitions

directed network

alternative bipartite representation

matching in original network

—— matched link   —— unmatched link   ● matched node   ● unmatched node (driver)
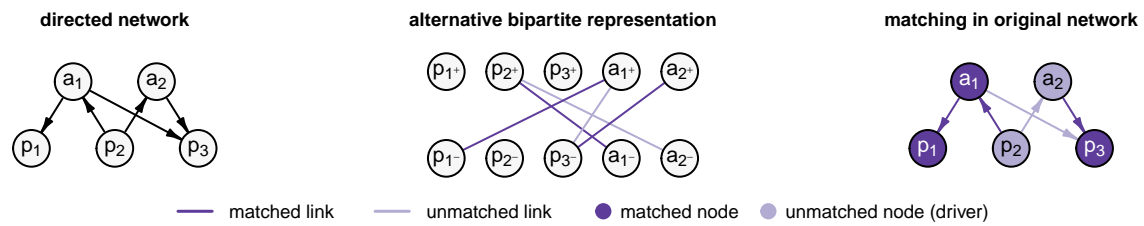
Figure S1: Finding a maximum matching in a complex network. (left) Directed network that indicate the direction of control between species. (center) Alternative bipartite representations of the directed networks. (right) The matchings in the bipartite representation mapped back to the original network.

pollination network

directed network

a maximum matching

—— matched link   —— unmatched link   ● matched node   ● unmatched node (driver)

a	*Geum urbanum*	**f**	*Epistrophe grassulariae*	**k**	*Meligethes aeneus*
b	*Heracleum sphondylium*	**g**	*Episyrphus balteatus*	**l**	*Miridae*
c	*Rubus fruticosus*	**h**	*Fannia sp.*	**m**	*Orthotylus/Lygocorus*
d	*Botanophila fugax*	**i**	*Lasioglossum rufitarse*	**n**	*Parasitica*
e	*Sciapus sp.*	**j**	*Melanostoma scalare*	**o**	*Byturus tomentosus*

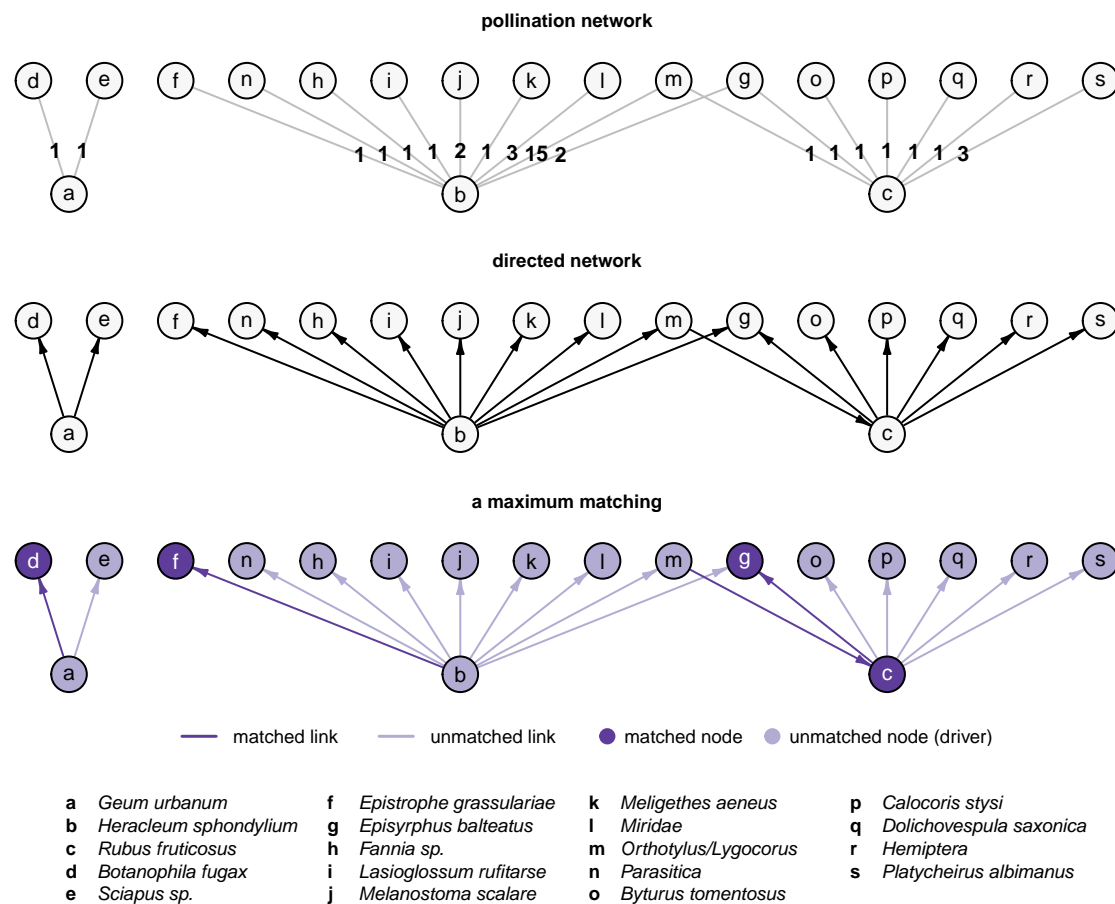|	
p	*Calocoris stysi*
q	*Dolichovespula saxonica*
r	*Hemiptera*
s	*Platycheirus albimanus*

Figure S2: Illustration of the procedure with an empirical community. The visitation network (top), the number of visits between species pairs are shown on each link. The directed network in which the direction of control is determined based on the mutual dependences (middle). One of the possible maximum matchings of this network (bottom).

P: alternative bipartite representation   **L(P): line graph of P**   **H: complement of L(P)**   **max. cliques in H**
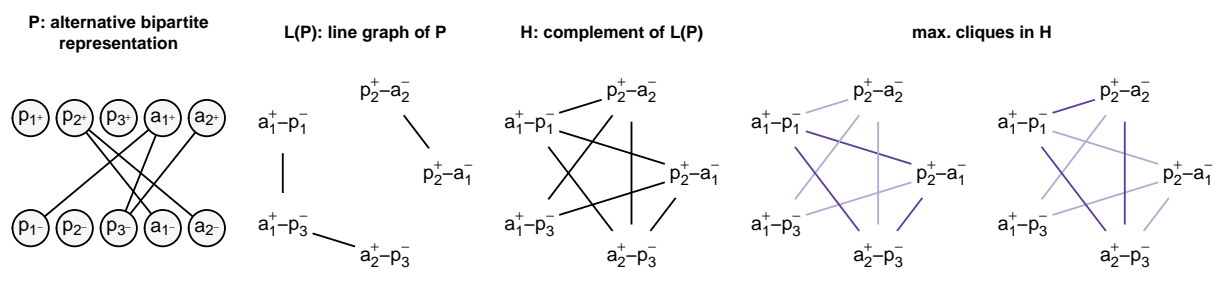

Figure S3: Finding all possible maximum matchings. From left to right: alternative bipartite representation of the directed network. Line graph of the network. The complement of the network. The two maximal cliques are shown in dark purple.

are necessary. First, H is a graph with the same nodes as $L(P)$ but that has a link between two nodes if and only if there is not a link in $L(P)$. Second, a clique is a subset of nodes such that all pairs of them are linked. Lastly, a maximal clique is a clique such that there are no cliques composed of more nodes (Gutin, 2013). In this example, there are two maximal cliques. The final step is then to map these cliques onto the original network to obtain all possible maximal cardinality matchings as shown in Figure 2 in the main text.
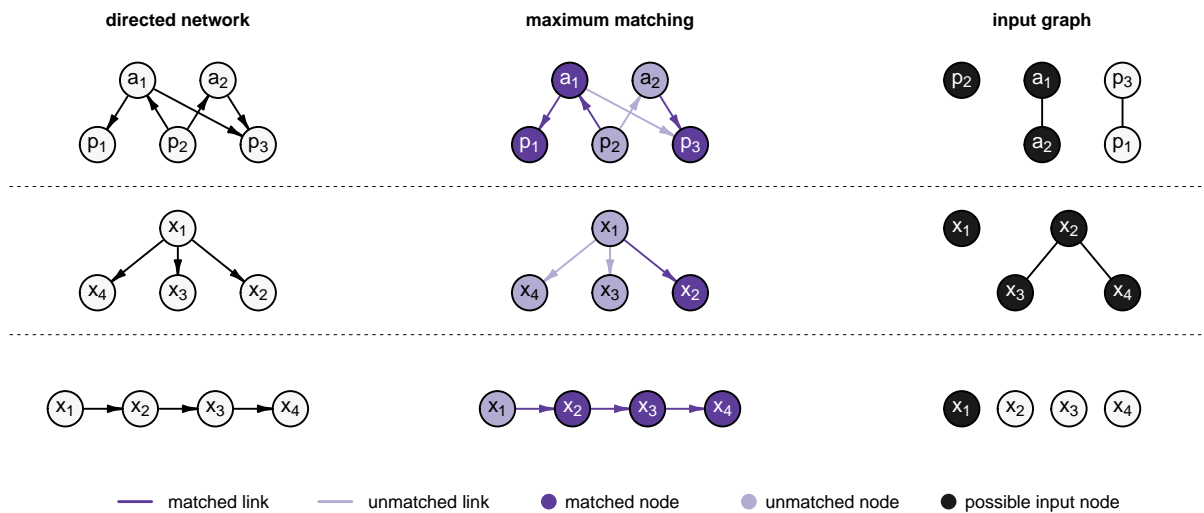
Figure S4: Input graph. Two nodes x_i and x_j are said to be control adjacent if there is a node x_k connecting x_i and x_j with an unmatched link $x_k - x_i$ and a matched link $x_k - x_j$. For example, in the top panel, a_1 is control adjacent to a_2 (via p_2), p_1 is control adjacent to p_3 (via a_1) and p_2 has no control adjacent nodes. As a_1 and p_2 belong to the minimum driver node set under the control configuration found with the maximum matching shown, then every node in the same component as these two nodes is also a possible driver node. By combining the nodes from the components with minimum driver node sets is possible to enumerate the different control configurations. For example the network in the top panel has two possible minimum driver node sets $[p_2, a_1]$ and $[p_2, a_2]$ (Figure 2). The network in the middle panel has three minimum driver node sets $[x_1, x_2]$, $[x_1, x_3]$, and $[x_1, x_4]$. Finally the network in the bottom pannel has only one $[x_1]$.

S3.3 Control adjacency and the input graph

The algorithm proposed by Zhang, Lv, & Pu (2016) enable us to find all minimum driver node sets without calculating all possible maximum matchings. The algorithm is based on the construction of the input graph of the directed network. Components in the input graph reveal the correlations between nodes from a structural control perspective (Figure S4). The input graph is constructed based on the *control adjacency* of nodes in a maximum matching. Two nodes x_i and x_j are said to be control adjacent if there is a node x_k connecting x_i and x_j with an unmatched link $x_k - x_i$ and a matched link $x_k - x_j$. For example a_1 is control adjacent to a_2 in the top panel of Figure S4. If a node x_i is part of a minimum driver node set then every node in the same component of x_i in the input graph is a possible driver node.

In order to calculate a node's control capacity ϕ, we harness the fact that nodes can be classified into three groups according to their role in the input graph. First, if x_i belongs to the minimum driver node set and has no adjacent nodes, then it cannot be replaced and will be part of every possible control configuration. Second, if x_i does not belong to the minimum driver node set and is not adjacent to any possible input node, then it does not part of the minimum driver node set under any control configuration. Finally, if a node x_i belongs to the minimum driver node set and is adjacent to another node x_j in the input graph, then x_j must belong to another minimum driver node to which x_i does not belong, in other words, x_i and x_j are substitutable.

The first and the second group have a control capacity of $\phi = 1$ and $\phi = 0$ respectively. Meanwhile, the third group has intermediate values of control capacity which can be calculated by computing the number of possible substitutions of a node by its control adjacent nodes. As a rule of thumb, the largest the number of possible substitutions an input node has, the smaller its control capacity.

For larger networks, calculating the control capacities using the input graph is several orders of magnitude faster than computing all possible maximum matchings. While the computation time of the input graph grows linearly with network size, the computation of all maximum matchings grows polynomially. However, the control capacities obtained using the input graph are not identical to those obtained by computing all possible maximum matchings. The reason for this discrepancy is that two maximum matchings might result in the same minimum driver node set. Nevertheless, the results obtained using both approaches are extremely similar. We compared the control capacity obtained by calculating all maximum matchings and that obtained using the input graph in 13 of our networks (those for which we were able to calculate all maximum matchings within a reasonable time). The Spearman correlation was very high in all cases, ranging between 0.89 and 1 (median 0.98).
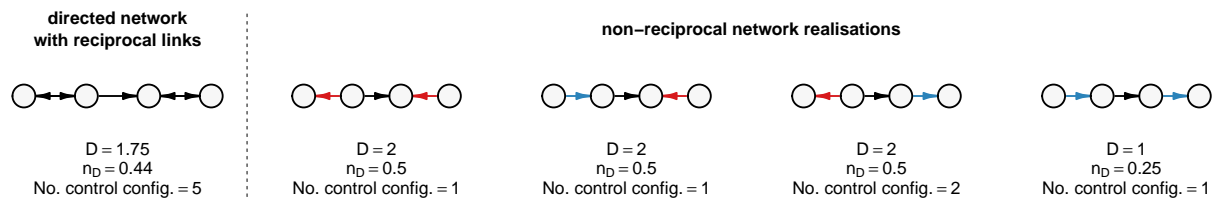
Figure S5: In order to calculate the controllability metrics for a network with reciprocal links (left), we first decompose the network into several versions that contain no reciprocal links (right). As each of these networks represents a different equiprobable scenario, the controllability metrics for the network with reciprocal links are the unweighted mean of these metrics across the non-reciprocal networks.

S3.4 Networks with reciprocal links

Our procedures to find minimum driver node sets are only adequate when there are no reciprocal (bidirectional) links in the network. The proportion of reciprocal links in relatively low. In fact, the proportion of species pairs that are reciprocally connected in our empirical networks is just 2.97%. However, reciprocal links are not a rare occurrence as 16 out of 20 empirical networks had at least one. To calculate our controllability metrics at both the network (the relative size of the minimum driver node set, n_D) and species level (control capacity, ϕ, and the likelihood of being a superior node, σ) we first need to generate all possible versions of the of the network that include only non-reciprocal links (Figure S5).

All of these non-reciprocal networks are as likely to be one that best represents the superior/matched relationships in the network. As such, we average the three metrics (n_D, ϕ, and σ) across all the networks without reciprocal links without weighting by the number of possible control configurations.

9

S4   Structural stability

We follow R. P. Rohr, Saavedra, & Bascompte (2014) and S. Saavedra, Rohr, Olesen, & Bascompte (2016) to calculate the structural stability of the empirical networks. We first calculate the stability condition $\hat{\gamma}$ using all species in the network. To calculate the contribution to stable coexistence of a given species, we removed the focal species from the network and then calculated the corresponding structural stability Ω for a level of mutualism equal to the stability condition found previously. For simplicity, in the main text we show results for when there is a mutualistic trade-off of $\delta = 0$ and a mean interspecific competition of $\rho = 0.01$ (S. Saavedra et al., 2016). The choice of the parameters had only a minor impact on the relative contribution of species to stable coexistence (Figure S6a) and overall there was no change on our result that indicates that critical species (those with $\phi = 1$) have a larger than average contribution to stable coexistence (Figure S6b).
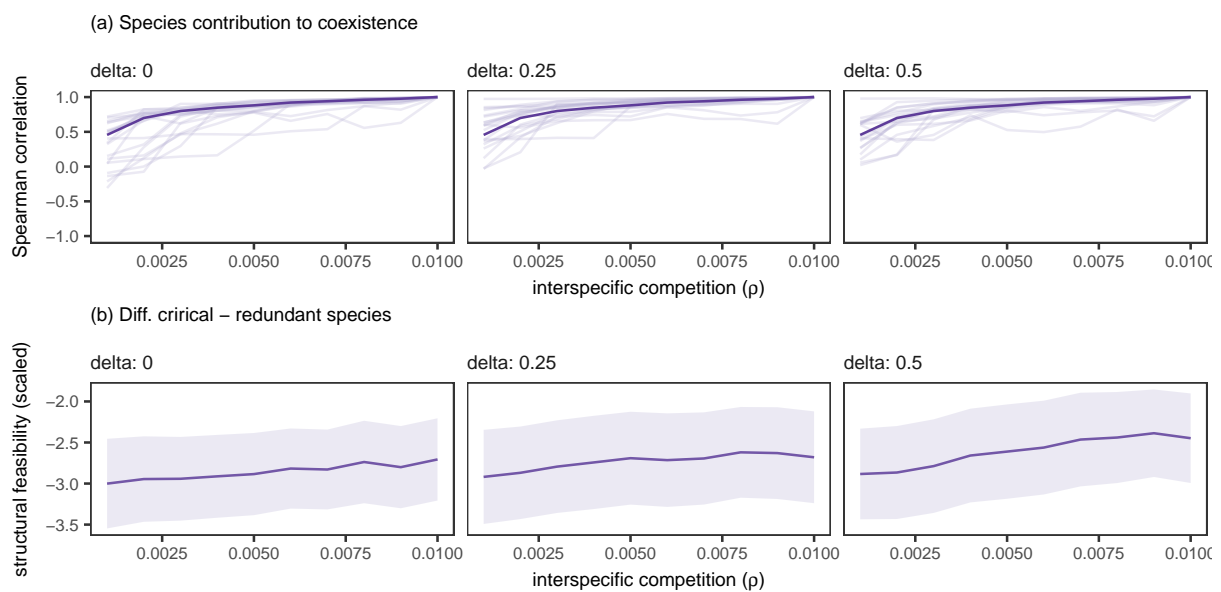
(a) Species contribution to coexistence

(b) Diff. crirical – redundant species

Figure S6: A sensitivity analysis of the impact of interspecific competition parameter ρ on structural stability reveals that the choice of parameters has a negligible impact on the results. (a) Spearman correlation coefficient of the contribution to stable coexistence of species. Lighter lines indicate the correlation coeficient for each of the twenty empirical networks. The dark line indicates the median value. (b) Difference between the conribution to stable coexistence of critical and redundant species for different values of the iterspecific competition parameters ρ.

S5   Visitation as a proxy for species dependence

Visitation frequency has been shown to be an appropriate surrogate for inter-specific effects in pollination networks (Bascompte, Jordano, & Olesen, 2006; Vázquez, Morris, & Jordano, 2005). Nevertheless, visitation is not equivalent to pollen deposition and might be insufficient to reflect the dependences of plants on animals and vice versa (Alarcón, 2010; King, Ballantyne, & Willmer, 2013). We, therefore, investigated the effect of calculating the dependences using visitation or pollination effectiveness and importance—two metrics more proximate to plant reproductive success. We did this by comparing *(i)* the size of the minimum driver node set n_D of the community, *(ii)* the species' control capacity and, *(iii)* the species' probability of being a superior node. To do this, we used data collected by G. Ballantyne, Baldock, & Willmer (2015) from a low diversity pollination community at a dry lowland heathland in Dorset, UK (50° 43.7'N 2° 07.2'W). First, deposition networks were quantified using the mean single visit deposition—the number of conspecific pollen grains effectively deposited on a virgin stigma during a single visit by a particular animal (G. Ballantyne et al., 2015; King et al., 2013; Ne'eman, Jürgens, Newstrom-Lloyd, Potts, & Dafni, 2009). Second, visitation networks were constructed by counting the visits to flowers during Single Visit Depositions. Finally, pollinator importance networks were constructed as the product of pollinator efficiency and visit frequency.

At a network scale the driver-node density n_D was consistent among the three weighting schemes (0.67 for deposition, 0.67 for the visitation, and 0.62 for the pollinator-importance network, respectively). The choice of weighting can also have an impact on at the species level. Therefore we calculated ϕ and σ and calculated it's correlation among all three weighting schemes. Although visitation and efficiency (pollen deposition) produce moderately different results, we found a very strong agreement between the order produced by visitation and importance which is arguably a more accurate metric of interspecific effects (Table S2).

Table S2: Spearman correlation coefficient matrix between the control capacity and the probability obtained by weighting links by the visitation, pollinator efficiency and pollinator importance.

	efficiency	importance	visitation
control capacity			
efficiency	1.00	0.37	0.38
importance	0.37	1.00	0.86
visitation	0.38	0.86	1.00
superior probability			
efficiency	1.00	0.69	0.57
importance	0.69	1.00	0.90
visitation	0.57	0.90	1.00

Altogether, the evidence supports the idea that visitation is a suitable metric to estimate the mutual dependence of species pairs. First, it is directly related to pollinator foraging. Second, it produces results within our controllability framework that are consistent with plant reproductive success (as estimated by

the importance metric).

13

S6　Sensitivity to sampling

Our approach is fundamentally based on the network structure. Often, the majority of the interactions that make up this structure are weak; in our networks, this means that most interactions are formed by a small number of observed pollination visits, and therefore those weak interactions are less conspicuous in the field than strong ones. To strengthen the case of our approach, we, therefore, evaluated the robustness of our results to simulated sampling limitations.

To do so we removed a portion of the visits, for each network, and calculated how three control metrics of the subsampled network compare to those of the full network. Specifically, we calculated the difference between the relative size of the minimum driver node set n_D of the subsampled and the full network, the Spearman correlation between the control capacities ϕ and the probability of being a superior node σ that were obtained for the empirical network and subsamples of it. Finally, we also compared the control capacity of critical species ($\phi = 1$) across the different subsamples. We removed up to 20% of the interactions in 1% increments. We repeated the procedure for each network to obtain a total sample size of $n = 400$.

Overall, the results of the sensitivity analysis indicate that our approach is likely to still be useful in the absence of complete sampling. Specifically, as the proportion of sampled interaction decreased, the variability of n_D increased but was overall similar to that obtained when using all available interactions (Figure S7). Similarly, the correlation between ϕ and σ of empirical and subsampled networks was high even for extreme levels of undersampling (Figure S8). Importantly, critical species in the empirical networks were also critical species in the subsampled networks in a large majority of cases (Figure S9).
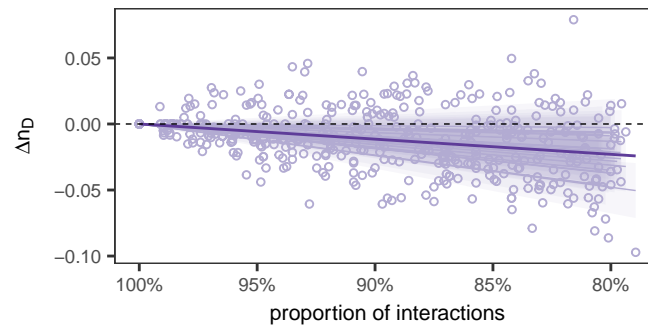
Figure S7: Difference between the size of the minimum driver node set of the empirical network and the subsampled network. Lighter lines indicate the changes for each networks and the darker line indicates the overall trend.
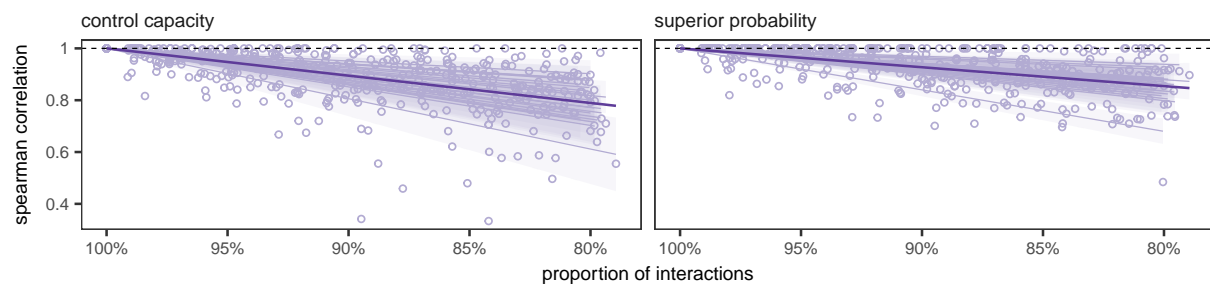

Figure S8: Spearman correlation between the control capacity and the probability of being a superior nodes of the empirical networks and subsampled networks. Lighter lines indicate the changes for each networks and the darker line indicates the overall trend.
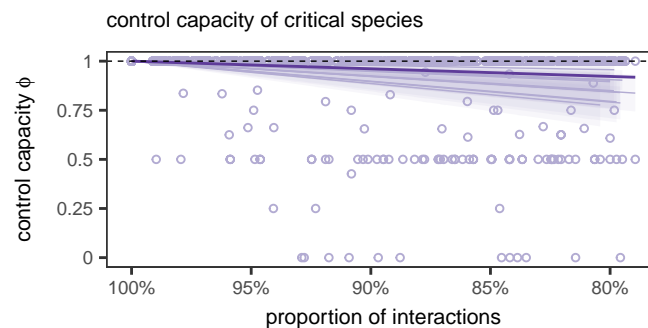

Figure S9: Control capacity of critical species (those with $\phi = 1$ in the empirical network) in subsampled networks. Lighter lines indicate the changes for each networks and the darker line indicates the overall trend.

Table S3: Model selection table based on the Akaike Information Criterion for small samples for candidate models of size of the minimum driver node set (n_D) in the empirical networks. There was no evidence for any of the evaluated predictors: the invasion status (inv), the network connectance (c), the network nestedness (NODF) the number of species (n_S) the network asymmetry (AS_s) or the interaction strength asymmetry (AS_s). Only candidate models with a weight > 0.01 are shown.

(Intercept)	C	AS_s	inv	n_s	study	AS_n	NODF	df	delta	weight
0.89	-0.40	—	—	-0.20	—	-0.50	-0.19	6	0.00	0.21
1.07	—	-0.11	—	—	—	-0.31	—	4	1.33	0.11
0.95	-0.38	—	—	-0.20	—	-0.50	—	5	1.46	0.10
1.03	—	-0.10	—	—	—	-0.31	-0.15	5	2.23	0.07
0.84	-0.42	—	—	-0.22	+	-0.46	-0.20	7	2.91	0.05
1.02	-0.11	-0.13	—	—	—	-0.31	—	5	2.91	0.05
0.98	-0.11	-0.12	—	—	—	-0.31	-0.16	6	3.66	0.03
1.07	—	—	—	—	—	-0.38	—	3	3.79	0.03
1.02	—	—	—	—	—	-0.38	-0.18	4	3.90	0.03
0.91	-0.34	-0.04	—	-0.16	—	-0.45	-0.18	7	4.09	0.03
0.97	-0.30	-0.06	—	-0.14	—	-0.43	—	6	4.32	0.02
0.90	-0.40	—	—	-0.21	+	-0.47	—	6	4.43	0.02
0.88	-0.41	—	+	-0.21	—	-0.51	-0.20	7	4.65	0.02
1.06	—	-0.12	—	0.02	—	-0.29	—	5	4.75	0.02
1.07	—	-0.11	+	—	—	-0.31	—	5	4.93	0.02
1.07	—	-0.11	—	—	+	-0.31	—	5	4.95	0.02
0.95	-0.38	—	+	-0.20	—	-0.51	—	6	5.62	0.01
1.08	—	—	—	-0.04	—	-0.40	—	4	5.97	0.01

Table S4: Summary table of the control capacity model with the smallest AICc.

term	estimate	std.error	statistic	p.value
(Intercept)	-1.20	0.37	-3.21	0.00
visitation strength	0.29	0.16	1.86	0.06
contribution to nestedness	0.15	0.14	1.05	0.29
asymmetry	0.80	0.25	3.14	0.00
level (pol.)	2.70	0.39	6.96	0.00
contribution to nestedness:level (pol.)	0.32	0.19	1.70	0.09
asymmetry:level (pol.)	-1.98	0.28	-7.17	0.00

Table S5: Variable importance for control capacity models. The importance is the sum of the weights of all models in which the variable was included.

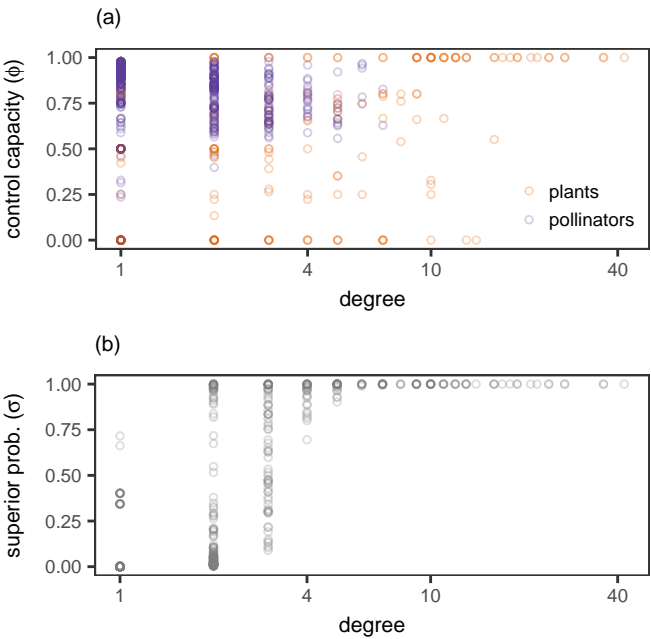| term	importance
level (pol.)	1.00
asymmetry	1.00
asymmetry:level (pol.)	1.00
visitation strength	0.93
contribution to nestedness	0.91
contribution to nestedness:level (pol.)	0.79
visitation strength:level (pol.)	0.39
degree	0.28
degree:level (pol.)	0.25

Figure S10: Relationship between degree and (a) control capacity and (b) the probability of being a superior node.

Figure S11: Mean Spearman correlation coefficients between control metrics (control capcity ϕ and the likelyhood of being a superior node σ) and centrality metrics commonly used to estimate a species' 'keystoness' (closeness centrality l , betwenness w, eigen centrality e, page rank g and degre k) in each of the networks. Purple and orange tiles indicate a positive correlation or negative correlation respectively.

References

Alarcón, R. (2010). Congruence between visitation and pollen-transport networks in a California plant-Pollinator community. *Oikos*, *119*(1), 35–44. doi:10.1111/j.1600-0706.2009.17694.x

Almeida-Neto, M., & Ulrich, W. (2011). A straightforward computational approach for measuring nestedness using quantitative matrices. *Environmental Modelling & Software*, *26*(2), 173–178. doi:10.1016/j.envsoft.2010.08.003

Ballantyne, G., Baldock, K. C. R., & Willmer, P. G. (2015). Constructing more informative plantPollinator networks: Visitation and pollen deposition networks in a heathland plant community. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1814), 20151130. doi:10.1098/rspb.2015.1130

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

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, *1695*(5), 1–9.

Gutin, G. (2013). Independence and Cliques. In P. Zhang (Ed.), *Handbook of Graph Theory* (Second Edi, pp. 475–489). Chapman and Hall/CRC. doi:10.1201/b16132-31

King, C., Ballantyne, G., & Willmer, P. G. (2013). Why flower visitation is a poor proxy for pollination: Measuring single-visit pollen deposition, with implications for pollination networks and conservation. *Methods in Ecology and Evolution*, *4*(9), 811–818. doi:10.1111/2041-210X.12074

Liu, Y.-Y., & Barabási, A.-L. (2016). Control principles of complex systems. *Reviews of Modern Physics*, *88*(3). doi:10.1103/RevModPhys.88.035006

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

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

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

Saavedra, S., Rohr, R. P., Olesen, J. M., & Bascompte, J. (2016). Nested species interactions

promote feasibility over stability during the assembly of a pollinator community. *Ecology and Evolution*, *6*(4), 997–1007. doi:10.1002/ece3.1930

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

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