Quantifying the manageability of pollination networks in an invasion context

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

[1] Centre for Integrative Ecology, School of Biological Sciences, University of Canterbury, Private Bag 4800, Christchurch 8041, New Zealand

[2] Current address: Department of Ecology, Swedish University of Agricultural Sciences, Box 7044, SE-750 07 Uppsala, Sweden

Author for correspondence: Daniel Stouffer (daniel.stouffer@canterbury.ac.nz) - +64 3 364 2729 - Centre for Integrative Ecology, School of Biological Sciences, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

Abstract

Here, we leverage recent advances in control theory to assess the "manageability" of ten pairs of uninvaded and invaded plant-pollinator communities. We found that the networks' manageability is most strongly determined by the ratio of plants to pollinator richness, which in turn constrains the networks' degree distribution. We also characterise species' suitability for inclusion in management interventions by exploring the entire space of control strategies. Specifically, we measure the extent to which species (i) are necessary to steer the state of the community and (ii) are able to affect the abundance of other species. We found that invasive plants have a dominant position in every invaded community and that this dominance is underpinned by high asymmetry in the dependence of their interaction partners. Our results highlight the advantages of the control-theoretic framework for ecological questions and provide novel insight into the design of informed management interventions.

Keywords: Network control theory, ecological networks, species importance, driver species, mutualism, management interventions

Introduction

In a complex system, the whole is often greater than the sum of its parts [1–3]. Within community ecology, a complex-systems approach has led to the development of analytical and simulation tools with which to understand, for example, the role of species embedded in a network of interactions [4–6]. The inherent complexity of nature, however, has regularly hindered—or at least complicated—our ability to find management solutions to many problems ecological communities face. To overcome this obstacle, we require a framework that allows us to explain, predict, and manage ecological communities, particularly when they are confronted with perturbations [7,8]. Ideally, such a framework is equipped to account for their complex structure and the dynamics that determine the species abundances and the state of the community.

Among the various possibilities, *control theory* appears to be a strong candidate [9]. Widely used in engineering to determine and supervise the behaviour of dynamical systems [10], it is well equipped to deal with the many feedbacks present in ecological communities [11]. Research in this area has established a strong link between the structure of complex networks and their *controllability*—the relative ability to manipulate network components to drive the system to a desired state [12–14]. These advances suggest that it is in principle possible to alter a whole ecological community's composition by modifying the abundances of only a few species. Applications of control theory to ecological networks can also take into account the extent to which changes in the abundances of one species may ripple through the community [13]. Therefore, control theory could also be harnessed to help identify which species are most relevant from a structural and dynamic perspective.

This information is valuable not only for basic ecology, but it might be also relevant to address more applied management and conservation challenges. This is particularly true in the context of biotic invasions, where identifying key players in the community is a prerequisite to informed attempts to alter the state of invaded ecosystems and maintain the state of uninvaded ones. Despite recent advances in network theory, practical challenges to the conservation of interaction networks persist [15], and the link between the structure of complex networks and our ability to manage and conserve them is still ambiguous [16,17]. To complicate things further, biotic invasions can induce dramatic changes in the patterns of interactions that determine the structure of ecological networks [18–20], in particular pollination [21–25]. Understanding how the differences in network structure before and after invasion impact our ability to manage the communities is thus a double challenge, but it is also the critical first step towards a fully informed recovery. Despite the apparent overlap, the control-theoretic perspective has not been adopted in an invasion context.

To bridge this gap, we outline an approach to apply control theory in an ecological context and implement it using empirical data. Specifically, we use a set of ten pairs of uninvaded and invaded plant-pollinator communities to investigate the link between invasion, network structure and ecological management. While doing so, we focus on two particular questions. First, grounded in the difficulties usually involved with invasive-species eradication and ecosystem restoration [26], we ask whether invaded networks have lower levels of "manageability" than their uninvaded counterparts; that is, whether they require a greater proportion of species to be managed to achieve the same level of control. Second, we ask whether some species are more important than others at driving the population dynamics of the community and which factors determine this importance.

Theoretical framework

At the core of representing ecological communities as complex systems sits the idea that the state of any given node depends on its state, the state of the nodes it interacts with, the state of the nodes they interact with, and so on. These dependencies are given by the structure of the interactions, the dynamic relationships between nodes, and the mechanisms of self-regulation. This representation has been very useful for our understanding of ecological communities and might be advantageous for a more formal approach to ecological management based on control theory.

The overall objective of control theory is to be able to steer a system from one state to another in finite time. However, in complex systems like ecological communities, the large number of nodes and interactions as well as their non-linear dynamics render its control extremely challenging. On one theoretical extreme, for instance, any ecological community could be fully controlled if we control the state of every species independently. On another, the mechanisms of self-regulation and the multiple feedback cycles found in ecological communities mean it is also theoretically possible to control the whole community by directly modifying the state of just a *single* species [27,28]. Although mathematically correct, neither of these options are practical in real-world applications: the first because it is infeasible to design and implement interventions that modify the abundance of every single species in a community, and the second because the control signal might require unreasonably large amounts of time, energy, or unattainable rapid changes on the state [10,29]. Here we explore an intermediate point between these two trivial solutions.

In particular, we leverage the principle of species interdependence to find an intermediate set of *driver nodes* to control the network [11]. When we focus on net interspecific effects, it is possible to identify a minimum number of nodes the control of which can theoretically drive the state of every other node in the network to a desired configuration. Conveniently the information necessary to determine this minimum number of driver nodes D is fully contained in the network structure [10,12,30]. Such a system can be described by $\frac{dx}{dt} = \mathbf{A}x + \mathbf{B}u(t)$, 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 the species respond to the external input, respectively. If S is the number of species in the community, the matrix $\mathbf{A}$ has size $S \times S$ whereas the matrix $\mathbf{B}$ has size $S \times D$.

The goal of structural controllability, which we employ here, is to use the information contained in $\mathbf{A}$ to generate a supportable estimate of $\mathbf{B}$ (and by extension D). 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 our 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 [11,12].

Manageability

The number of driver nodes D provides a structural indication of how difficult a network's control might be. This is because systems that require a large number of external input signals are intuitively more difficult or costly to control. In an ecological context, external inputs that modify the state of a node can be thought of as management interventions. Therefore, the density of driver nodes $n_D = \frac{D}{S}$, where S is the total number of species in the community, is a measure of the extent to which network structure can be harnessed for network control. For instance, a hypothetical "network" in which species do not interact would require direct interventions for every single species to achieve full control, whereas a linear food chain would require just one species to be directly controlled to harness cascading effects through its trophic levels. From this perspective, it is possible to use n_D as an index of the manageability of an ecological community, understood in the context of how difficult is to modify the abundances of species in the community using external interventions—a common theme in ecosystem management, conservation, and restoration.

It has been recently shown that calculating D is equivalent to finding a maximum matching in the network [12]. In a directed network, a *matching* consists of a subset of links in which no two of them share a common starting or ending node (Figure 1, Supporting Information S1). A given matching has *maximal cardinality* if the number of matched links (also referred to as the *matching size*) is the largest possible. A maximal cardinality matching is then called a *maximum matching* if the sum of the weights of the matched links (also referred to as *matching weight*) is again the largest possible [31].

Once we have the subset of links that constitute a matching, we can also classify the nodes in the network based on that matching (Figure 1). A node is called *matched* if it is at the end of a matched link and *unmatched* otherwise. A node is also called *superior* if it is at the start of a matched link. Note that a node cannot be superior if it has no outgoing links. Notably, these node categories are what helps us to link a maximum matching back to the concept of network controllability, as follows. Unmatched nodes are the *driver nodes* D because they have no superior in the network and must be directly controlled by an external input [12]. Each matched node, on the other hand, can be controlled by its superior.

Note that this framework requires a directed network in which the direction of the links corresponds to the direction of control. In the "Methods" section below, we explain our approach to determining the link direction in pollination networks.

Relative importance

While calculating n_D measures the manageability of an ecological community, it does not provide information about the identity of the species that compose the set of driver nodes. Ecologically, potential differences between species are relevant because management and conservation resources are limited, and therefore ecological interventions should be focused on the set of species that might provide the largest impact. Moreover, maximum matchings in a network are often not unique, and each maximum matching indicates unique paths that can potentially be used to control the network. We harness this property and use a network's complete set of maximum matchings to characterise each species' relative importance in driving the state of the community. One possibility is to characterise a species by the frequency f_D with which it is classified as a driver node within this set of matchings. However, the profile of our networks indicates that a large proportion of species were classified as driver nodes because of external interventions required to achieve **full** controllability and not because they influence the abundance of other species (Supporting Information S2). Furthermore, the precise role of driver nodes is more ambiguous when full control is unfeasible or undesired—often the case in ecological settings. We therefore also calculate the frequency f_S with which a species is classified as a superior node since this is the frequency with which they form part of possible control paths.

Most commonly, structural controllability assumes unweighted networks—links exist or not, and hence f_D and f_S can be calculated by computing all possible maximum-cardinality matchings. However, we take the link weights into account when calculating the matchings here because the weights can reveal significant ecological patterns and processes that might be undetectable in unweighted networks [32–35]. Additionally, species A may interact with both species B and C but depends strongly on B and only weakly on C. Intuitively, a management intervention designed to indirectly modify the abundance of species A is more likely to succeed if the abundance of B, rather than C, is directly controlled. A complication of including the interaction weight when calculating the maximum matching, however, is that empirical interaction strengths are to some extent stochastic and depend on proximate factors such as sampling method and intensity [36]. We overcome this issue by calculating all maximal cardinality matchings and then ranking them by their matching weight. By following this approach, we effectively give priority to the species that participate in the pathways that potentially have the largest impact on the community while acknowledging the limitations associated with sampling and its potential restrictions [37].

Methods

We next applied the previously defined framework to ten paired pollination networks. Each network pair was composed of a community invaded by a plant and a community "free" of the invasive species. Four pairs were obtained from natural or semi-natural vegetation communities in the city of Bristol, UK [38]. These networks are comprised of 19–87 species (mean 55), and non-invaded plots were obtained by experimentally removing all the flowers of the invasive species *Impatients grandulifera*. The other six pairs were obtained from lower diversity Mediterranean shrublands in Cap de Creus National Park, Spain [23]. These networks are comprised of 30–57 species (mean 38); in contrast to the above, uninvaded communities were obtained from plots that had not yet been colonised by either of the invasive species *Carpobrotus affine acinaciformis* or *Opuntia stricta*. Further details about the empirical networks can be found in Supporting Information S3.

We then specified the structure of all networks using pollinator visitation frequency, which has been shown to be an appropriate surrogate for interspecific effects in pollination networks [39,40]. To further examine whether this decision would influence our results, we also evaluated the effect of using pollinator efficiency or pollinator importance as alternative measures in a different data set that lacked invasive species [41,42], and we found quantitatively similar results for all three options (Supporting Information S4). Because our approach depends on the network topology, we also evaluated the robustness of our results to the undersampling of interactions. Specifically, we calculated n_D and species relative importance for 500 random subsamples of each empirical network in which the weakest links were more likely to be removed. Our sensitivity analysis indicated that, even in the absence of complete sampling, a control-theoretic approach can still be applied (Supporting Information S5).

Manageability

We began by quantifying the manageability of each network. To do so, we calculated the networks' maximum matching and determined the minimum proportion of species n_D that need external input signals to fully control the species abundances in the community. Note that because all maximum matchings have the same matching size, it is only necessary to calculate one of them. To simplify the analysis, if a network had more than one component (two species are in different components if there exists no path between them and are hence independent of each other in terms of network control) we only considered the largest. Smaller components were present in eleven out of the twenty networks and were typically composed of just one plant and one pollinator. Their removal represented an average loss of 4.7% of the species and 2.7% of the interactions.

Weighting & directing links

As we noted earlier, the maximum-matching approach requires a directed network in which a link from species i to species j indicates that the abundance of j can be affected by the abundance of i. This implies that we need first to identify a directionality for the links between species that is consistent with the dynamics of the community (Figure 2). In some ecological networks, establishing the directionality can appear relatively straightforward, for example when links represent biomass transfer or energy flow [9]. Interspecific effects in pollination networks, however, are not strictly directed since the benefit is mutual between interacting species. Nevertheless, the relative extent to which a given pair of interacting species affect each other can be quantified by the magnitude of the mutual dependence between them [40]. The dependence of plant i on pollinator j, d_{ij}, is the proportion of the visits from pollinator j compared to all pollinator visits to plant i. Likewise, the dependence of pollinator j on plant i, d_{ji}, is the proportion of the visits by pollinator j to plant i compared to all visits of pollinator j. Using dependences generates a weighted bipartite network in which all interacting pairs are connected by two directed links (Figure 2b).

Given the respective dependences, the extent to which species i affects species j relative to the extent to which j affects i can be summarised by the interaction asymmetry [40] given by

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

Previous research has shown that mutual dependences are often highly asymmetric in natural communities [40]; in other words, if a plant species is largely dependent on a pollinator species, then that pollinator tends to depend rather weakly on the plant (or vice versa). We therefore simplified the network so that interacting species are connected by only one directed link when mutual dependences are asymmetric (Figure 2c). This simplification, while maintaining ecological realism, is advantageous for several reasons. First, it is consistent with previous advances in structural controllability; second, it prevents the singular case in which the network could be perfectly matched rendering $D = 1$ as well as problems related to the introduction of artificial control cycles [14]; and third it significantly reduces the computational resources necessary for the application of our approach (Supporting Information S6). Moreover, we found that changing to unidirectional interactions based on the direction of asymmetry does not alter the relative n_D of different networks (Table S3).

To find a maximum matching in a network with interaction directions and weights determined by the asymmetry, we adopted a strategy based on an alternative bipartite representation of the directed network with two levels that indicate the outgoing and incoming links to each node (Supporting Information S1). Once we had this alternative representation we used the maximum bipartite matching algorithm implemented in the `max_bipartite_match` function of igraph 1.0.1 [43] on each network.

Statistical analysis

We also wanted to test whether invasion status or other predictors had an impact on the observed values of n_D. We therefore used a set of generalised linear models (with binomial error structure) to investigate the effect of invasion status while also including covariates related to species richness, since one might naively expect to see a negative relationship between richness and manageability [44]. These covariates included the total number of species, plant richness, pollinators richness, the ratio of plant to pollinator richness, the link density (connectance), and the study site (as a two-level factor).

We next explored whether real networks differ in their architecture from random ones in a concerted way that affects manageability. Previous research indicates a direct link between a network's degree distribution and the number of nodes necessary to fully control it [12], but the strength and applicability of this relationship have not been tested for in weighted ecological networks. We therefore compared the driver-node density n_D of the empirical networks to networks generated by a null model that maintained each species' strength (its total sum of visits) while allowing their degrees (its number of interactions) to vary. Beyond network structure, the dependence asymmetry plays a fundamental role in determining the direction of control in each two-species interaction and therefore has the potential to influence the network n_D results above. We therefore performed an additional randomisation in which we kept the structure of each network constant but randomised the direction of the interaction asymmetries. That is, we first calculated the observed asymmetries for each community and then shuffled the direction of the link between each pair of species.

Additional details about the statistical models and the randomisations can be found in Supporting Information S7.

Relative importance

Our second key question was related to how species differ in their ability to drive the population dynamics of the community. To quantify this importance, we computed all maximal cardinality matchings in each network. We then calculated the frequency with which each species i was a driver (f_D) or a superior node (f_S) in the set of matchings that had a matching weight greater or equal to 0.8 times the weight of the maximum matching. We selected this threshold as it provided a high agreement between networks quantified by visitation and pollination efficiency as well as between our weighting/directionality assumptions; however, the choice of this threshold had a negligible impact on any results (Supporting Information S8). Details about the computational procedure to find all maximal cardinality matchings of a network can be found in Supporting Information S1 and Figure S2.

Statistical analysis

We then examined whether any species-level structural properties could predict our metrics of species importance—the frequency with which a species was a driver or a superior node (f_D and f_S, respectively). We used a set of generalised linear mixed-effects models (with binomial error structure) with the relative frequencies as the response variables. As predictors in this model, we included measures of centrality (degree and eigen-centrality), which have been found to be strong predictors of importance in a coextinction context [45]; a measure related to network robustness (contribution to nestedness), as nestedness has been proposed as one of the key properties that promote stability in mutualistic networks [46]; a measure of strength of association (visitation strength, the sum of visits a species receives or performs) and a measure of strength of dependence (species strength, the sum of dependences of all species on the focal species), as their distribution determines the extent of interspecific effects [40]. In addition, we also included guild (plant or pollinator) and whether the species is invasive or not as categorical fixed effects. Lastly, we allowed for variation between different communities by including the network identity as a random effect [47]. Supplementary details about the statistical models can be found in Supporting Information S7.

Results

Manageability

All of the networks had a driver-node density n_D between 0.55 and 0.92 (mean 0.76; Figure 3a). In addition, we found that, when controlling for potential species richness effects, the n_D of invaded communities was smaller to those of non-invaded communities (Figure 3b). Nevertheless, of the various covariates we explored, the ratio of plants to pollinators showed the strongest relationship with n_D (Figure 3c; Table S5). Specifically as the proportion of pollinators increases and the ratio plant/pollinator approaches unity, n_D decreases (all our communities had more pollinators than plants). Other covariates—connectance (link density) and species richness—had negative, but comparatively less important, relationships with n_D.

When exploring the effect of network structure itself, we observed that the driver-node density n_D of empirical networks was, in general, not significantly different to the manageability of network randomisations that maintained the degree of individual species (Figure 4). However, we found that the n_D of empirical networks was significantly larger than that of randomisations that maintained the network structure but that differed only in the direction of the asymmetries.

Relative importance

Invasive species were classified as superior nodes *and* driver nodes in every single network they were present; that is, they always had the highest relative f_S and f_D (Figure 5a). The model results suggest that these differences between invasive and native species are not underpinned by any intrinsic property of the invasive species; instead, they are due to species properties that apply to invasive and native species alike (Table 1). Specifically, we found that a species is more likely to be classified as a superior node if it had a large species strength (the sum of the dependences of all other species on the species of interest). To a smaller extent, visitation strength (a species' sum of visits) and degree also had a positive relationship with f_S. In contrast, the relationships between species structural properties and f_D were less clear cut (Table 1). Both invasive species (which generally have a larger degree and high dependence strength) and pollinators (which generally have a smaller degree and low dependence strength) were classified as driver nodes in a large proportion of matchings.

Discussion

Contrary to our initial hypothesis, we found some evidence that invaded communities might be easier to manage than uninvaded ones from a control-theoretic perspective. Our results reveal, however, that this effect is comparatively small, and the structural differences among different networks are more strongly related to potential differences in our ability to alter the state of the community via external interventions. Despite the small effect of invasion status at the network level, we found that invasive mutualists occupy a particularly dominant role in their communities for two reasons. First, as species with a high f_S, changes on their abundance have the potential to propagate broadly through the community and, in turn, affect the abundances of many other species. Second, as species with a high f_D, they are also indispensable when it comes to fully controlling the plant-pollinator network. At a community level, we demonstrate that the manageability of mutualistic networks is strongly governed by the asymmetric nature of mutual dependences—which constitute the foundations of the structure and stability of mutualistic networks [40,45,48–50]. Moreover, these mutual dependences seem to be constrained by the effects of both the patterns of species richness at each trophic guild and a network's degree distribution [51,52]. Indeed, the difference between the driver-node density (n_D) of our empirical networks and that of randomisations depended strongly on the null model's randomisation approach. While the empirical n_D was indistinguishable from that of networks with a random structure that maintained the degree of each species in the community, it was larger than that of randomisations in which the directed network was unchanged but where the observed patterns of dependence were broken.

Invasive species have been previously found to exacerbate the asymmetries in their communities [22,23,53]. Although this might cause differences both at the community and the species level, we found that invasive plants are not inherently different to their native counterparts [54,55]. Invasive plants, just like any other mutualist in our data set, tend to be classified as a superior node proportional to the degree to which their interaction partners are collectively more dependent on them than the other way around. Previous studies have found that supergeneralists, like invasive species, play a central role in their networks [24,56]. Our results take this one step further and indicate that dependence strength, rather than generalism or other metrics of centrality, is the factor that best explains the cascading effects a species could trigger on its community.

Because of the ability that our approach has to infer the magnitude of the effects that each species has on others in the community, it is tantalising to use it to select promising candidates for management interventions. To this end, the two indices we have used to characterise a species provide two complementary pieces of information. Our first index f_D—the frequency with which a species is classified as a driver node—provides an indication of the likelihood that a species forms part of the set of species that must be manipulated in order to control *all* species in the community. This driver-node concept has received considerable attention in the structural-control literature and indeed shows substantial potential to provide useful ecological insight. Nevertheless, we anticipate two caveats that hinder its direct utility for management applications. First, unlike some other types of complex systems, fully controlling an ecological community is almost certainly out of reach for all but the simplest, due to either the number of required interventions or the practical difficulties of their implementation [10]. For instance, our results suggest that full control of the pollination networks would require direct interventions on anywhere from 40–90% of the species. Second, [14] established that driver nodes arise due to distinct mechanisms and therefore species with markedly different network metrics can act as driver nodes in their community (Supporting Information S2).

Our second index f_S, however, is directly related to the likelihood a species affects the abundance of another species in all of the control strategies considered. Importantly, this is irrespective of whether controlling the entire network is ultimately desired and/or feasible. In fact, because superior nodes are always at the beginning of a matched link, species with a high f_S are more likely to be the subjects of management interventions when controlling the abundances of a target set of species—as opposed to the entire network—is desired [57]. An important advantage here is that the target set of species does not have to be the same set to which interventions are applied. For instance, despite inconsistent outcomes in practice [58–60], our results suggest that current restoration approaches that focus on direct eradication of invasive species might indeed be an effective way to modify ecosystem state. Nevertheless, our results also indicate that removals must be exercised with caution. Not only it is hard to predict the direction in which the system will change, but invaded communities also tend to be highly dependent on invaders and therefore acutely vulnerable to their eradication [25,61].

Despite the apparent similarities, our approach is different to previous attempts to quantify species importance in a few key ways. Existing metrics usually harness species features, like centrality, position, co-extinction or uniqueness, to infer their effect on other species [62–65]. In contrast, our control-based approach tackles that question directly. Although they are relatively simple to calculate, classic species-level network metrics do not necessarily reveal the best set of species to manage [66,67]. Our approach, however, is not based on a single structural metric but instead acknowledges the existence of multiple management strategies. By allowing for the fact that some strategies are better than others depending on the context, control theory implicitly highlights that management decisions should not be based on a single technique. As such, ours and other flexible approaches that take a network-wide approach might prove more useful to guide ecosystem management [67].

In this study, we illustrate how a control-theoretic approach can be employed in network ecology to evaluate the effect of invasions and other kinds of perturbations. We study the different ways a management intervention can be structured and provide a starting point for the continued study of controllability in ecological contexts. Although our pollination-specific results might not be directly translatable to other ecological networks that do not have bipartite structures, the approach we propose is applicable wherever species abundances are influenced by their interactions, and exciting open questions lie ahead. How to design the precise "control signals" to reach a desired ecosystem state or conservation outcome? What are the implications of assuming fully nonlinear dynamics? How important it is to include several interaction types for our understanding of manageability and species importance? What are the implications for species coexistence? Which are the trade-offs between persistence at the species and the community level? Answering each of these questions might require its evaluation in different ecological systems, an explicit integration of control theory with numerical models of species densities [13,68], and experimental tests on simple communities. Nevertheless, the potential rewards are encouraging from both an ecological and conservation perspective, where an integrated approach can shift the focus beyond the identification of ideal targets for intervention to the design of informed interventions that legitimately achieve restoration goals.

Appendix 1: Glossary

Driver node An unmatched node in a maximal cardinality matching or a maximum matching. From the control perspective, driver nodes are those to which external control signals must be applied in order to gain full control of the network.

Matched/unmatched link A link is refered to as *matched* if it is part of a matching, and *unmatched* otherwise.

Matched/unmatched node A node is refered to as *matched* if it is at the end of a matched link, and *unmatched* otherwise.

Matching A set of links in which no two of them share a common starting or ending node.

Matching size The number of matched links in a matching.

Matching weight The sum of the weights of all matched links in a matching.

Maximal cardinality matching A matching with the largest possible matching size. In unweighted/binary networks, all maximal cardinality matchings are also maximum matchings.

Maximum matching A matching with the largest possible matching size *and* largest possible matching weight.

Superior node The node at the start of a matched link. From the control perspective, superior nodes make up the chains that propagate the control signals through the network.

Data accessibility

All data used in this manuscript have already been published by [38], [23], and [42]. The reader should refer to the original sources to access the data.

Competing interests

We have no competing interests.

Author contributions

DBS conceived the idea; all authors contributed to the development of the theoretical framework. EFC performed all analysis. EFC and DBS wrote the manuscript. All authors contributed to its revision.

Acknowledgments

The authors thank Jane Memmott and coauthors, and eveyone that has made their data available to us, Takeuki Uno for the insight provided to find the set of all maximum matching algorithms, and Jason Tylianakis, Bernat Bramon Mora, Guadalupe Peralta, Rogini Runghen, Warwick Allen, Matthew Hutchinson, and Marilia Gaiarsa for feedback and valuable discussions.

Funding

EFC acknowledges the support from the University of Canterbury Doctoral Scholarship, the University of Canterbury Meadow Mushrooms Postgraduate Scholarship, a New Zealand International Doctoral Research Scholarship, and a travel grant from the European Space Agency. DBS acknowledges the support of a Marsden Fast-Start grant and a Rutherford Discovery Fellowship, administered by the Royal Society of New Zealand.

References

1. Jørgensen SE, Mejer H, Nielsen SN. 1998 Ecosystem as self-organizing critical systems. *Ecological Modelling* **111**, 261–268. (doi:10.1016/S0304-3800(98)00104-5)

2. Levin S. 1999 *Fragile Dominion*. Cambridge, MA: Perseus.

3. Montoya M, Pimm SL, Sole RV. 2006 Ecological networks and their fragility. *Nature* **442**, 259–264. (doi:10.1038/nature04927)

4. Pascual M, Dunne JA, editors. 2005 *Ecological Networks: Linking Structure to Dynamics in Food Webs*. New York: Oxford University Press.

5. Bascompte J, Stouffer DB. 2009 The assembly and disassembly of ecological networks. *Philosophical Transactions of the Royal Society B: Biological Sciences* **364**, 1781–7. (doi:10.1098/rstb.2008.0226)

6. Stouffer DB, Sales-Pardo M, Sirer MI, Bascompte J. 2012 Evolutionary Conservation of Species' Roles

in Food Webs. *Science* **335**, 1489–1492. (doi:10.1126/science.1216556)

7. Solé RV, Montoya JM. 2001 Complexity and fragility in ecological networks. *Proceedings of the Royal Society B* **268**, 2039–45. (doi:10.1098/rspb.2001.1767)

8. Green JL *et al.* 2005 Complexity in Ecology and Conservation: Mathematical, Statistical, and Computational Challenges. *BioScience* **55**, 501–510. (doi:10.1641/0006-3568(2005)055[0501:CIEACM]2.0.CO;2)

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

10. Motter AE. 2015 Networkcontrology. *Chaos* **25**, 097621. (doi:10.1063/1.4931570)

11. Liu Y-Y, Barabási A-L. 2016 Control principles of complex systems. *Reviews of Modern Physics* **88**, 035006. (doi:10.1103/RevModPhys.88.035006)

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

13. Cornelius SP, Kath WL, Motter AE. 2013 Realistic control of network dynamics. *Nature Communications* **4**, 1942. (doi:10.1038/ncomms2939)

14. Ruths J, Ruths D. 2014 Control profiles of complex networks. *Science* **343**, 1373–6. (doi:10.1126/science.1242063)

15. Tylianakis JM, Laliberté E, Nielsen A, Bascompte J. 2010 Conservation of species interaction networks. *Biological Conservation* **143**, 2270–2279. (doi:10.1016/j.biocon.2009.12.004)

16. Blüthgen N. 2010 Why network analysis is often disconnected from community ecology: A critique and an ecologist's guide. *Basic and Applied Ecology* **11**, 185–195. (doi:10.1016/j.baae.2010.01.001)

17. Kaiser-Bunbury CN, Blüthgen N. 2015 Integrating network ecology with applied conservation: a synthesis and guide to implementation. *AoB Plants* **7**, plv076. (doi:10.1093/aobpla/plv076)

18. Baxter CV, Fausch KD, Murakami M, Chapman PL. 2004 Fish invasion restructures stream and forest food webs by interrupting reciprocal prey subsidies. *Ecology* **85**, 2656–2663. (doi:10.1890/04-138)

19. Tylianakis JM, Didham RK, Bascompte J, Wardle DA. 2008 Global change and species interactions in terrestrial ecosystems. *Ecology Letters* **11**, 1351–1363. (doi:10.1111/j.1461-0248.2008.01250.x)

20. Ehrenfeld JG. 2010 Ecosystem Consequences of Biological Invasions. *Annual Review of Ecology, Evolution, and Systematics* **41**, 59–80. (doi:10.1146/annurev-ecolsys-102209-144650)

21. Olesen JM, Eskildsen LI, Venkatasamy S. 2002 Invasion of pollination networks on oceanic islands: Importance of invader complexes and endemic super generalists. *Diversity and Distributions* **8**, 181–192.

(doi:10.1046/j.1472-4642.2002.00148.x)

22. Aizen MA, Morales CL, Morales JM. 2008 Invasive mutualists erode native pollination webs. *PLOS Biology* **6**, 0396–0403. (doi:10.1371/journal.pbio.0060031)

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

24. Vilà M, Bartomeus I, Dietzsch AC, Petanidou T, Steffan-Dewenter I, Stout JC, Tscheulin T. 2009 Invasive plant integration into native plant-pollinator networks across Europe. *Proceedings of the Royal Society B* **276**, 3887–3893. (doi:10.1098/rspb.2009.1076)

25. Traveset A, Heleno R, Chamorro S, Vargas P, Mcmullen CK, Nogales M, Herrera HW, Olesen JM, Mcmullen K. 2013 Invaders of pollination networks in the Galápagos Islands : emergence of novel communities. *Proceedings of the Royal Society B* **280**, 20123040. (doi:10.1098/rspb.2012.3040)

26. Woodford DJ, Richardson DM, Macisaac HJ, Mandrak NE, Van Wilgen BW, Wilson JRU, Weyl OLF. 2016 Confronting the wicked problem of managing biological invasions. *NeoBiota* **86**, 63–86. (doi:10.3897/neobiota.31.10038)

27. Cowan NJ, Chastain EJ, Vilhena DA, Freudenberg JS, Bergstrom CT. 2012 Nodal dynamics, not degree distributions, determine the structural controllability of complex networks. *PLoS ONE* **7**. (doi:10.1371/journal.pone.0038398)

28. Rahmani A, Ji M, Mesbahi M, Egerstedt M. 2009 Controllability of Multi-Agent Systems from a Graph-Theoretic Perspective. *SIAM Journal on Control and Optimization* **48**, 162–186. (doi:10.1137/060674909)

29. Yan G, Ren J, Lai YC, Lai CH, Li B. 2012 Controlling complex networks: How much energy is needed? *Physical Review Letters* **108**, 1–5. (doi:10.1103/PhysRevLett.108.218703)

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

31. West D. 2001 *Introduction to graph theory*. Upper Saddle River: Prentice Hall.

32. Scotti M, Podani J, Jordán F. 2007 Weighting, scale dependence and indirect effects in ecological networks: A comparative study. *Ecological Complexity* **4**, 148–159. (doi:10.1016/j.ecocom.2007.05.002)

33. Tylianakis JM, Tscharntke T, Lewis OT. 2007 Habitat modification alters the structure of tropical host-parasitoid food webs. *Nature* **445**, 202–205. (doi:10.1038/nature05429)

34. Vázquez DP, Melián CJ, Williams NM, Blüthgen N, Krasnov BR, Poulin R. 2007 Species abundance and asymmetric interaction strength in ecological networks. *Oikos* **116**, 1120–1127. (doi:10.1111/j.2007.0030-

1299.15828.x)

35. Kaiser-Bunbury CN, Muff S, Memmott J, Müller CB, Caflisch A. 2010 The robustness of pollination networks to the loss of species and interactions: a quantitative approach incorporating pollinator behaviour. *Ecology Letters* **13**, 442–52. (doi:10.1111/j.1461-0248.2009.01437.x)

36. Gibson RH, Knott B, Eberlein T, Memmott J. 2011 Sampling method influences the structure of plant-pollinator networks. *Oikos* **120**, 822–831.

37. Jordano P. 2016 Sampling networks of ecological interactions. *Functional Ecology* **30**, 1883–1893. (doi:10.1111/1365-2435.12763)

38. Lopezaraiza-Mikel ME, Hayes RB, Whalley MR, Memmott J. 2007 The impact of an alien plant on a native plant-pollinator network: An experimental approach. *Ecology Letters* **10**, 539–550. (doi:10.1111/j.1461-0248.2007.01055.x)

39. Vázquez DP, Morris WF, Jordano P. 2005 Interaction frequency as a surrogate for the total effect of animal mutualists on plants. *Ecology Letters* **8**, 1088–1094. (doi:10.1111/j.1461-0248.2005.00810.x)

40. Bascompte J, Jordano P, Olesen JM. 2006 Asymetric Coevolutionary Networks Facilitate Biodiversity Maintenance. *Science* **312**, 431–433. (doi:10.1126/science.1123412)

41. Ne'Eman G, Jürgens A, Newstrom-Lloyd L, Potts SG, Dafni A. 2010 A framework for comparing pollinator performance: Effectiveness and efficiency. *Biological Reviews* **85**, 435–451. (doi:10.1111/j.1469-185X.2009.00108.x)

42. Ballantyne G, Baldock KCR, Willmer PG. 2015 Constructing more informative plant-pollinator networks: visitation and pollen deposition networks in a heathland plant community. *Proceedings of the Royal Society B* **282**, 20151130. (doi:10.1098/rspb.2015.1130)

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

44. Menge BA. 1995 Indirect effects in marine rocky intertidal interaction webs: patterns and importance. *Ecological Monographs* **65**, 21–74. (doi:10.2307/2937158)

45. Memmott J, Waser NM, Price MV. 2004 Tolerance of pollination networks to species extinctions. *Proceedings of the Royal Society B* **271**, 2605–2611. (doi:10.1098/rspb.2004.2909)

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

47. Bates D, Mächler M, Bolker B, Walker S. 2015 Fitting Linear Mixed-Effects Models using lme4.

Journal of Statistical Software **67**, 1–48. (doi:10.18637/jss.v067.i01)

48. Vázquez DP, Aizen MA. 2004 Asymmetric specialization: a pervasive feature of plant-pollinator interactions. *Ecology* **85**, 1251–1257. (doi:10.1890/03-3112)

49. Lever JJ, Nes EH van, Scheffer M, Bascompte J. 2014 The sudden collapse of pollinator communities. *Ecology Letters* **17**, 350–359. (doi:10.1111/ele.12236)

50. Astegiano J, Massol F, Vidal MM, Cheptou P-O, Guimarães PR. 2015 The Robustness of Plant-Pollinator Assemblages: Linking Plant Interaction Patterns and Sensitivity to Pollinator Loss. *PLOS ONE* **10**, e0117243. (doi:10.1371/journal.pone.0117243)

51. Melián CJ, Bascompte J. 2002 Complex networks: two ways to be robust? *Ecology Letters* **5**, 705–708. (doi:10.1046/j.1461-0248.2002.00386.x)

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

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

54. Stouffer DB, Cirtwill AR, Bascompte J. 2014 How exotic plants integrate into pollination networks. *Journal of Ecology* **102**, 1442–1450. (doi:10.1111/1365-2745.12310)

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

56. Palacio RD, Valderrama-Ardila C, Kattan GH. 2016 Generalist Species Have a Central Role In a Highly Diverse Plant-Frugivore Network. *Biotropica* **48**, 349–355. (doi:10.1111/btp.12290)

57. Gao J, Liu Y-Y, D'Souza RM, Barabási A-L. 2014 Target control of complex networks. *Nature Communications* **5**. (doi:10.1038/ncomms6415)

58. Suding KN, Gross KL, Houseman GR. 2004 Alternative states and positive feedbacks in restoration ecology. *Trends in Ecology and Evolution* **19**, 46–53. (doi:10.1016/j.tree.2003.10.005)

59. Rodewald AD, Rohr RP, Fortuna MA, Bascompte J. 2015 Does removal of invasives restore ecological networks? An experimental approach. *Biological Invasions* **17**, 2139. (doi:10.1007/s10530-015-0866-7)

60. Smith AM, Reinhardt Adams C, Wiese C, Wilson SB. 2016 Re-vegetation with native species does not control the invasive Ruellia simplex in a floodplain forest in Florida, USA. *Applied Vegetation Science* **19**, 20–30. (doi:10.1111/avsc.12188)

61. Albrecht M, Padrón B, Bartomeus I, Traveset A. 2014 Consequences of plant invasions on compart-

mentalization and species' roles in plant-pollinator networks. *Philosophical Transactions of the Royal Society B: Biological Sciences* **281**, 20140773. (doi:10.1098/rspb.2014.0773)

62. Allesina S, Bodini A. 2004 Who dominates whom in the ecosystem? Energy flow bottlenecks and cascading extinctions. *Journal of Theoretical Biology* **230**, 351–358. (doi:10.1016/j.jtbi.2004.05.009)

63. Jordán F, Liu W-c, Davis AJ. 2006 Topological keystone species: measures of positional importance in food webs. *Oikos* **112**, 535–546. (doi:10.1111/j.0030-1299.2006.13724.x)

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

65. Lai S-M, Liu W-C, Jordan F. 2012 On the centrality and uniqueness of species from the network perspective. *Biology Letters* **8**, 570–573. (doi:10.1098/rsbl.2011.1167)

66. Eklöf A, Tang S, Allesina S. 2013 Secondary extinctions in food webs: A Bayesian network approach. *Methods in Ecology and Evolution* **4**, 760–770. (doi:10.1111/2041-210X.12062)

67. McDonald-Madden E, Sabbadin R, Game ET, Baxter PWJ, Chadès I, Possingham HP. 2016 Using food-web theory to conserve ecosystems. *Nature Communications* **7**, 10245. (doi:10.1038/ncomms10245)

68. Gibson TE, Bashan A, Cao H-t, Weiss ST, Liu Y-y. 2016 On the Origins and Control of Community Types in the Human Microbiome. *PLOS Computational Biology* **12**, e1004688. (doi:10.1371/journal.pcbi.1004688)

Table legends

1 **Factors explaining species importance**. Factor estimates correspond to the average over all models that accounted for 95% of the AIC evidence. Confidence intervals correspond to $\alpha = 0.05$.

Figure legends

1 **Matchings of a simple network**. (a) We start with a network in which the direction of the links indicates the potential direction of control; for example a link from a_1 to p_1 indicates that the state of p_1 is influenced by the state of a_1. The numbers indicate the weight of each link. (b & c) This network has two maximal cardinality matchings; that is, two configurations in which it would be possible to exert full control of the network via external input signals to a minimal set of nodes. In both cases, the three matched links (purple arrows) represent the control paths through the network and provide an indication of the matched nodes (purple), which are controlled by superior nodes within the network (circular nodes). Unmatched nodes (orange) are called driver nodes because full network control requires external signals to be applied to them. Out of the two maximal cardinality matchings only one (c) has maximum weight and therefore is also a maximum matching. Further examples can be found in Supporting Information S1.

2 **Different ways to depict quantitative mutualistic networks**. (a) Pollination networks are frequently described by the observed number of visits between each plant and animal species. (b) Based on that visitation data, the mutual dependences between interacting species are calculated directly based on the relative visitation frequencies. (c) The relative differences of these dependences—the interaction asymmetry—then provide a means to estimate the dominant direction of the interspecific effects.

3 **Driver-node density**. (a) Histogram of the driver-node density (n_D) for the twenty networks. (b) Invaded communities have lower n_D than uninvaded communities even when controlling for factors related to species richness. The boxes cover the 25th–75th percentiles, the middle lines mark the median, and the maximum length of the whiskers is 1.5 times the interquartile range. (c) Out of the richness metrics, the ratio of plants to pollinators showed a strong, negative relationship with n_D. In both plots, partial residuals correspond to the partial working residuals of the invasion status in our generalised linear mixed model.

4 **The driver-node density n_D of empirical networks compared to network randomisations**. For each randomisation approach, we show the standarised rank of the empirical value compared to the set of randomisations. A scaled mean rank—akin to a p-value— less than 0.025 or greater than 0.975 (the areas shaded in light grey) suggests a significant difference between the empirical network and its randomisations. The empirical n_D is much larger than that of network randomisations in which the direction of asymmetries has been randomised. In contrast, the manageability of networks in which the species degrees were randomly shuffled were not significantly different. All boxes are as in Figure 3a.

5 **Relationships between f_S and f_D and species structural properties**. (a) In all networks were they were present, invasive species were classified as superior (f_S) and driver (f_D) nodes in all possible control configurations. (b) Species strength (the sum of the dependences of other species on the species of interest) is the single most important factor explaining f_S. Visitation strength and degree also had an important albeit comparatively smaller effect (dashed lines correspond to $\pm$ one standard deviation of these factors). Invasive species are depicted as circles.

Table 1

	imp.	est.	C.I.
f_s			
(Intercept)	1.00	2.69	2.5
species strength	1.00	34.26	15
visitation strength	1.00	1.37	1.1
degree	0.90	4.12	5.5
contribution to nestedness	0.56	0.44	1.3
guild (pollinator)	0.48	0.72	2.6
eigen-centrality	0.25	0.00	0.19
invasive sp.	0.24	-6.23	3.2E+06
f_d			
(Intercept)	1.00	-0.19	0.83
guild (pollinator)	1.00	4.05	0.99
contribution to nestedness	1.00	1.41	0.62
degree	1.00	-5.31	2.5
species strength	1.00	4.65	2.6
visitation strength	1.00	3.07	2.7
eigen-centrality	0.71	0.72	1.5
invasive sp.	0.08	10.95	4.5E+06

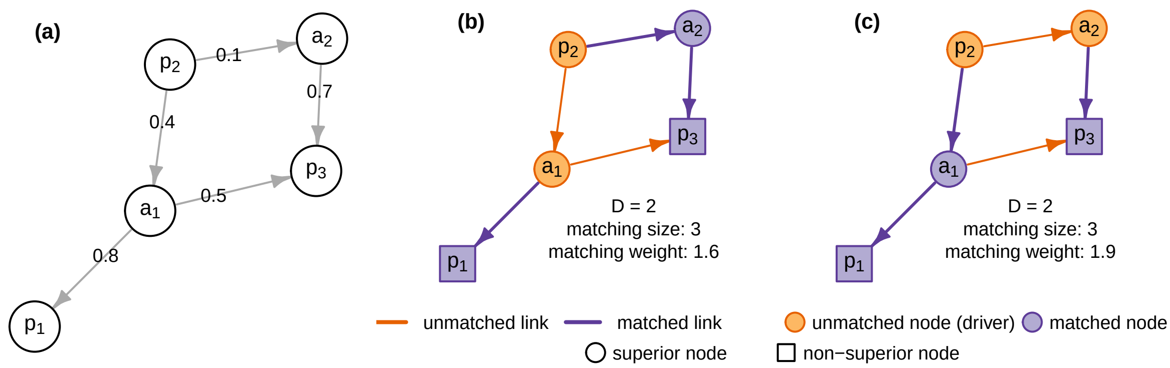

(a)
p2
a2
0.1
0.7
0.4
p3
0.5
a1
0.8
p1
(b)
p2
a2
p3
a1
D = 2
matching size: 3
matching weight: 1.6
p1
(c)
p2
a2
p3
a1
D = 2
matching size: 3
matching weight: 1.9
p1
unmatched link
matched link
unmatched node (driver)
matched node
superior node
non−superior node

Figure 1

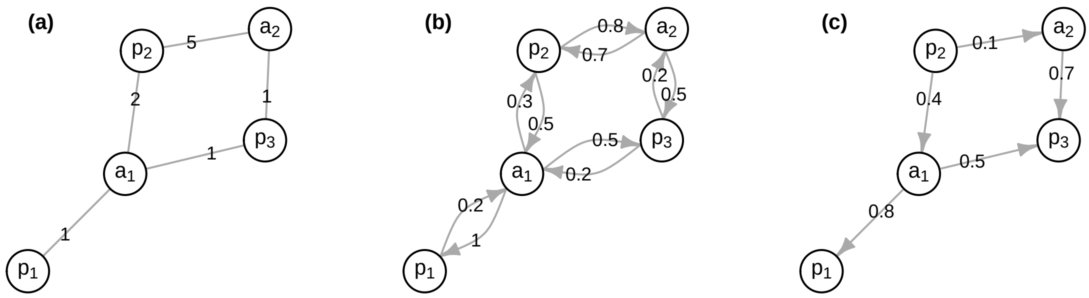
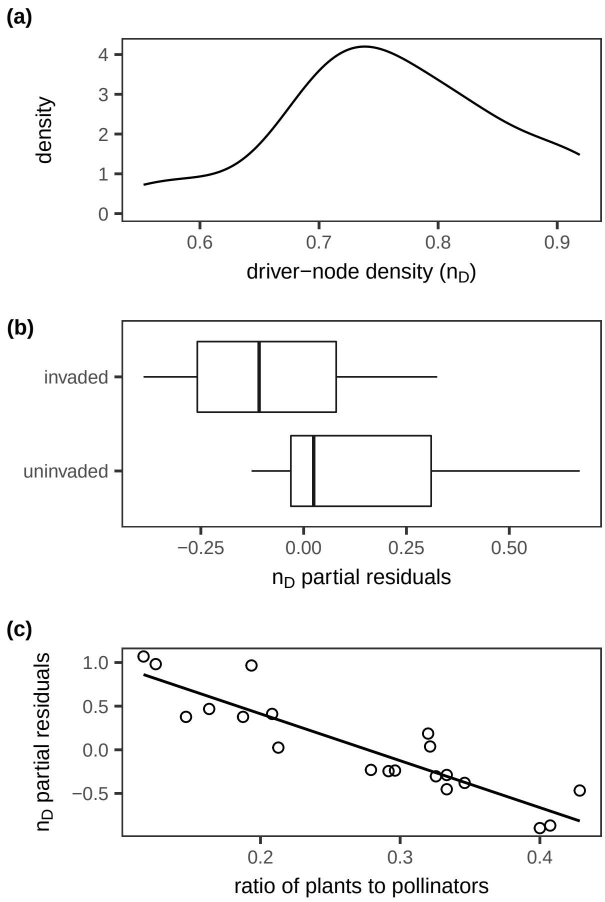

(a)
p2
a2
5
2
1
p3
1
a1
1
p1
(b)
p2
0.8
a2
0.7
0.2
0.3
0.5
0.5
0.5
p3
a1
0.2
0.2
1
p1
(c)
p2
0.1
a2
0.7
0.4
p3
a1
0.5
0.8
p1

Figure 2



(a)
density
4
3
2
1
0
0.6
0.7
0.8
0.9
driver−node density (nD)
(b)
invaded
uninvaded
−0.25
0.00
0.25
0.50
nD partial residuals
(c)
nD partial residuals
1.0
0.5
0.0
−0.5
0.2
0.3
0.4
ratio of plants to pollinators

Figure 3

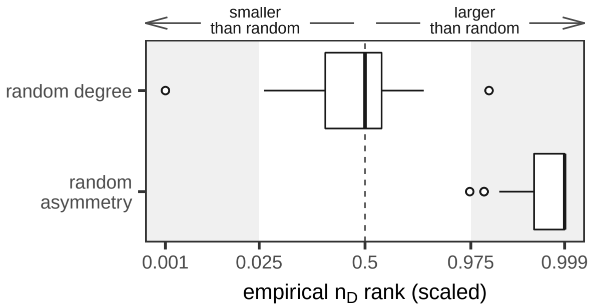

smaller
than random
larger
than random
random degree
random
asymmetry
0.001
0.025
0.5
0.975
0.999
empirical n_D rank (scaled)

Figure 4

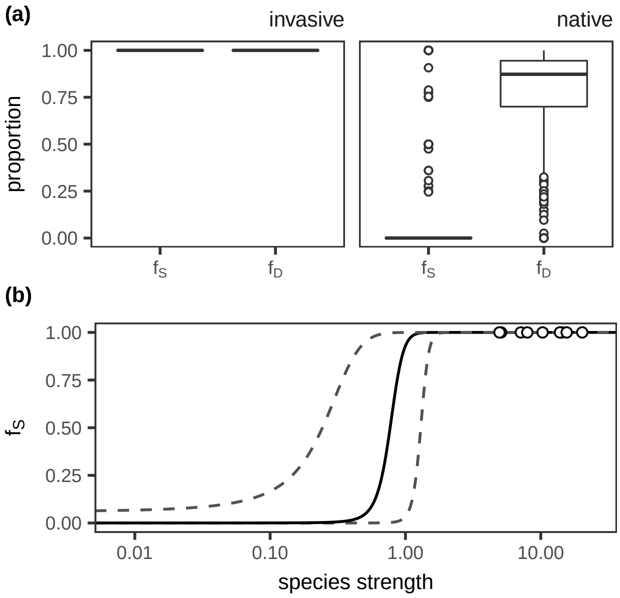

(a)
invasive
native
proportion
1.00
0.75
0.50
0.25
0.00
f_S
f_D
f_S
f_D
(b)
f_S
1.00
0.75
0.50
0.25
0.00
0.01
0.10
1.00
10.00
species strength

Figure 5