Quantization of ecological interactions yields insights into food web assembly and dynamics

Justin D. Yeakel,[1] Mathias Pires, James O'Donnell, Marcus de
Aguiar, Paulo Guimarães Jr, Dominique Gravel, and Thilo Gross

[1]*School of Natural Sciences, University of California, Merced, Merced, CA 95340, USA*

The dynamics of community assembly has a rich history in ecological theory. Recently, there has been much interest in assessing how changes in diversity within assembling communities impact the structure of interactions and vice versa. Here we examine a novel theoretical framework that seeks to generate communities by distilling many types of ecological interactions into a small number of unique pairwise directed links between species, including, but not limited to, 'assimilate' interactions (e.g. resource dependencies) and 'need' interactions (e.g. reproductive or habitat dependencies). Different pairwise combinations of directed link types between species give rise to the larger diversity of species interactions observed in nature, such as consumer-resource, and both service-resource and service-service mutualisms. Moreover, our framework permits the explicit inclusion of interactions that create or modify abiotic elements, which other species may depend on for survival, such as nutrients or habitat. Inclusion of both biotic and abiotic agents permits more complex and indirect interdependencies between species, effectively incorporating the concept of ecosystem engineering into interaction networks, where the environment can be altered by a species or groups of species, thereby facilitating or inhibiting the colonization of others. Our framework makes specific predictions that are borne out by observations in the field. First, we find that communities initially exhibit higher connectance (link density), which is quickly eroded to empirically observed values as competition for resources increases exclusion. Niche overlap among species follows similar trends: there is initially a greater degree of overlap between species, and as the system settles to a steady state, this overlap becomes minimized, mirroring observations of assembly in natural grasslands. Importantly, we find that an increase in the number of engineering species, by creating a greater number of direct and indirect interdependencies, constrains the assembly of communities initially, yet promotes assembly as the system matures. This leads to communities that exhibit greater diversity, however the increased species richness facilitated by engineers comes at a cost: as the number of engineers grows early in the assembly process, extinction cascades become larger. Our framework shows that despite the complexity of real communities, some of the most remarkable processes and patterns such as competitive exclusion, resource complementarity and extinction cascades, can be generated by simple interaction rules. Moreover our findings indicate that ecosystem engineering might be an important component that is overlooked in ecological network theory.

Introduction

What do we gain by focusing on community dynamics rather than population dynamics? By assuming that all species present in a community have achieved a steady state, we focus instead on those factors that

Model Description

The ENIgMa Model We seek to understand the compound nature of species interactions by describing them by a limited set of directional interactions that are combined to represent an ecological relationship. We aim to examine how these interdependencies between species in communities either aid or inhibit both assembly and extinction over long timescales, and specifically how ecosystem engineers contribute to these dynamics. We approach these questions by considering multiple types of directional interactions between species, which, when paired represent specific ecological relationships including trophic interactions, service-resource and service-service mutualisms, and commensalisms. We also introduce two types of nodes in our depiction of ecological networks: species and objects. Objects are made by species (here and henceforth refereed to as engineers) and represent a modification to the available niche-space for other species in the community, including but not limited to: an introduced compound, metabolite, or alteration to the habitat/environment.

The ENIgMa model consists of four directed interactions: e: eat, which specifies a dependency involving the exchange of biomass, n: need, which specifies a dependency that does not involve biomass flow (e.g. a reproductive service), i: ignore, the null interaction, and m: make, which connects a species to an object that it engineers. 'Objects' are interactive components that can be made by ≥ 1 species, and eaten, needed, or ignored by the others. The four directed interaction types describe specific dependencies that one species/object has on another, however it is the coupling of two opposing directed interactions that describe familiar ecological relationships (Table 1).

The e $\leftrightarrow$ i interaction describes a typical predator-prey relationship, where species 1 eats species 2, whereas aside from serving as a resource, species 2 does not interact with (ignores) species 1. Of course, a prey's abun-
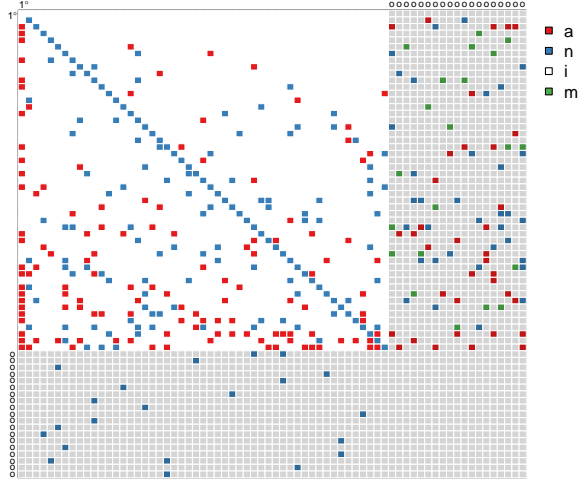
Figure 1: An example of the source pool interaction matrix where $\mathcal{S} = 50$. Species and objects are aligned across the rows and columns; objects are shaded and labeled by 'o' to distinguish them from species. The interaction recorded in row i and column j describes the directed interaction from species/object i to species/object j. The first row/column represents the basal resource; species that assimilate the primary resource are capable of primary production. Species interact with other species and/or objects; objects only interact with their engineers by 'needing' them; objects do not interact with other objects.

dance does not *ignore* the effects of predation, however our framework operates at the scale of presence/absence rather than abundance, and we assume that if both species co-occur, they have positive population densities, such that the prey's state (presence/absence) ignores the predator. A second type of trophic interaction is described by e ↔ e, where consumption is symmetric, which could be do to changing roles over the lifetime of an individual. The e ↔ n and n ↔ n interactions describe service-resource and service-service mutualisms, respectively. In the case of the former, one species interacts by way of a trophic interaction, whereas the other is provided a non-trophic need, such is the case in a plant-pollinator relationship. Unique to models of ecological networks, the m ↔ n interaction describes ecosystem engineering, where a species makes an object, whereas the presence of the object 'needs' the presence of the species that makes it. Objects can be utilized by other species in the community, providing an indirect dependency that could be facilitated by multiple species (many engineers make the same object) and/or used by multiple species (many species eat or need the same object). As we will describe below, engineers modify the niche space available to the community on timescales that last longer than the species themselves.

We explore the assembly of a novel community that emerges from a species source pool, which is represented by a source pool interaction matrix where all eat, need,

ignore, make interactions are established between all species and objects. As such, a set of interactions for a particular species defines how it interacts with any other *a priori*, thereby establishing its potential interaction niche space. The source pool is used to seed a novel community, which arises as the result of colonization and extinction rules, the details of which we describe below.

Building the source pool The source pool interaction matrix $\boldsymbol{P}$ is generated by first setting the number of species in the pool $\mathcal{S}_{\boldsymbol{P}}$ and determining the number of objects $\mathcal{O}_{\boldsymbol{P}}$ that are made by ecosystem engineers. The resulting matrix is $\mathcal{N}_{\boldsymbol{P}} \times \mathcal{N}_{\boldsymbol{P}}$ where $\mathcal{N}_{\boldsymbol{P}} = \mathcal{S}_{\boldsymbol{P}} + \mathcal{O}_{\boldsymbol{P}}$, and is subdivided into four quadrants, only two of which play a role here: species-species interactions and species-object interactions (see Fig. 1). In each quadrant, the expected frequency of eat interactions $\mathrm{E}\{p_e\}$ and the expected frequency of need interactions $\mathrm{E}\{p_n\}$ are free parameters, as is the expected number of objects made per species $\mathrm{E}\{\mathcal{O}_i\} = \eta$. Here and throughout, we simplify this parameter space by assuming that the frequency of eat and need interactions for species-species (SS) interactions and species-object (SO) interactions scale as ϕ, such that $\mathrm{E}_{\mathrm{SS}}\{p_e\} = \phi \mathrm{E}_{\mathrm{SO}}\{p_e\}$ and $\mathrm{E}_{\mathrm{SS}}\{p_n\} = \phi \mathrm{E}_{\mathrm{SO}}\{p_n\}$. For each species, a set number of objects is drawn from $\mathrm{Poiss}(\eta)$, such that the expected proportion of species that are engineers (species that make objects) is $1 - e^{-\eta}$. If a particular object is randomly and independently drawn for a given engineer from a complete list of all possible objects, such that multiple species – with some probability – can make the same object, the expected number of objects is

$$\mathrm{E}\{\mathcal{O}_{\boldsymbol{P}}\} = \mathcal{S}_{\boldsymbol{P}}\eta\left(1 - \frac{1}{e}\right), \qquad (1)$$

where e is Euler's number. The frequency of $m \leftrightarrow n$ interactions is then calculated as

$$\mathrm{E}\{p_m\} = \frac{\eta}{\mathcal{S}_{\boldsymbol{P}}\left(1 + \eta - \frac{\eta}{e}\right)^2}. \qquad (2)$$

Finally the frequency of the ignore interaction is calculated as $\mathrm{E}_{\mathrm{SS}}\{p_i\} = 1 - \mathrm{E}_{\mathrm{SS}}\{p_e\} + \mathrm{E}_{\mathrm{SS}}\{p_n\}$ and $\mathrm{E}_{\mathrm{SO}}\{p_i\} = 1 - \mathrm{E}_{\mathrm{SO}}\{p_e\} + \mathrm{E}_{\mathrm{SO}}\{p_n\} + \mathrm{E}_{\mathrm{SO}}\{p_m\}$ for species-species and species-object interactions, respectively. Pairwise interaction probabilities between both species and objects are then calculated as shown in Table I. These pairwise interactions are assigned randomly between species and objects independently in both quadrant, such that the source pool matrix has no imbued structure apart from the number of species, the number of objects, and the frequency of each directional interaction type. Each source pool is provided a *basal resource* (the first row/column). A species with a single trophic interaction to this resource is identified as a primary producer (Fig. 1), however the basal resource does not have eat, need, or make interactions itself.

Colonization and Extinction Assembly of a species community is the result of both local colonization and extinction of species that are drawn from the source pool.

The realized interactions within the assembled community $\boldsymbol{A}$ are thus a subset of the potential interactions observed if every species were present (as recorded in the source pool $\boldsymbol{P}$). We determine the ability of a species to colonize a community as a function of two conditions: 1) the colonizing species must eat *at least one* species/object in the community, and 2) the colonizing species must satisfy all of its need interactions; if these conditions are both satisfied, colonization is possible. At each time-step, one potential colonizer that fulfills these conditions is selected at random and added to the community, as well as the objects that it makes if it is an engineer. Thus, in the first time-step, only species that consume the primary resource (row 1; figure 1) and do not have any 'need' interactions can initiate the assembly process.

Extinction occurs directly via competitive exclusion, or indirectly via the subsequent loss of a consumer's single resource or any of the species/objects it needs. Extinction due to competitive exclusion is determined by violation of a single condition: a species must be the strongest competitor for at least one of its food resources. In a given community, each species i has a competition strength σ_i that is compared to that for every species j that shares its resources. If σ_i is not the highest σ for at least one its resources, then species i is competitively excluded from the community along with all unique objects that it makes. The competition strength for species i, σ_i, increases as the sum of its potential need interactions, and decreases as both the sum of its potential eat interactions and the sum of its realized predators, such that

$$\sigma_i = \pi \sum_{j=1}^{\mathcal{N}_P} \mathrm{n}_{\boldsymbol{P}(i,j)} - \sqrt{2} \sum_{j=1}^{\mathcal{N}_P} \mathrm{e}_{\boldsymbol{P}(i,j)} - \sum_{i=1}^{\mathcal{N}_A} \mathrm{e}_{\boldsymbol{A}(i,j)}, \quad (3)$$

where the summations describe the number of need interactions, eat interactions, and predators, respectively, and $\mathcal{N}_A = \mathcal{S}_A + \mathcal{O}_A$ is the size of the assembled community $\boldsymbol{A}$ as the sum of the number of species and objects. The coefficients serve only to prevent the substitution of different interaction types. Why do we assume that mutualisms increase a species' comepetitive strength? Although mutualisms serve to tie the existence of one species to another, which increases its risk exposure, we assume that this dependency evolved as the consequence of a fitness advantage inherent in the interaction, providing a competitive edge. Conversely, we assume that specialists (species with fewer trophic interactions) are competitively superior to generalists (ref), and that as a species spends more energy avoiding predation, it spends less energy competing. Importantly, we emphasize that the role of mutualisms and trophic interactions in determining a species' competition strength is with respect to its *potential* interaction niche, and thus calculated from the source pool matrix $\boldsymbol{P}$, whereas its vulnerability to predation is determined by a species' predators within the assembled community $\boldsymbol{A}$. Therefore, the first two factors of σ_i (the influence of its need

and eat interactions) are independent of the community, whereas its vulnerability changes with community structure. We note that due to the threshold conditions for colonization, $\sum_{j=1}^{\mathcal{N}_P} \mathrm{n}_{\boldsymbol{P}(i,j)} = \sum_{j=1}^{\mathcal{S}_A} \mathrm{n}_{\boldsymbol{A}(i,j)}$.

We integrate these colonization and extinction rules to simulate community assembly over time using a Gillespie algorithm. At each time-step a single event is chosen at random to iterate the simulation forward, where possible event types include: 1) species colonization, 2) species extinction, and 3) object extinction. The likelihood of drawing each event increases with the number of potential colonizers (n_c) or the number of species (n_e) and objects (n_o) that meet the conditions required for extinction. The change in time dt at each step in the simulation is therefore dynamic, where $\mathrm{dt} = 1/\sum_i k_i(t)$ where $k_i(t)$ is the number of possible events of type i at time t. This algorithm allows, with some probability, that both species and objects to remain in the system after they are selected for extinction, however this probability declines as the rate of colonization (number of potential colonization events, n_c) decreases and the rate of extinction (number of potential extinction events, $n_e + n_o$) increases. To incorporate the idea that the timescale of objects in a system should be longer than the timescale of the engineers that made them, we bias the algorithm to select a colonization or species extinction event with weight τ. If $\pi_e = n_e/n$ and $\pi_o = n_o/n$, species extinction occurs with probability $\tau \pi_e/(\tau \pi_e + (1-\tau)n_o)$ vs. an object extinction event that occurs with probability $(1-\tau)n_o/(\tau n_e + (1-\tau)n_o)$: as τ increases, the likelihood that an object is eliminated from the system declines.

Results & Discussion

Diverse interactions without engineers Community assembly in the absence of engineers ($\eta = 0$) reveals the emergence of food web and mutualistic web properties consistent with observations of steady state and assembling empirical systems. Because only primary producers that do not have outgoing need interactions can colonize initially, a diverse base of these autotrophs typically constitutes the early assembly process. In order for communities to have > 1 pure autotroph, we do not consider competitive exclusion of the basal resource, such that all non-mutualistic pure autotrophs have the potential to coexist. Following the establishment of a suite of autotrophs, both mixotrophs and low trophic-level heterotrophs assemble into the community, and competitive exclusion begins to increase the extinction rate. It is with the establishment of higher-trophic level consumers that the rate of extinction begins to approximate the rate of colonization, such that community diversity begins fluctuating around a steady state. This community steady state increases as the number of mutualisms established in the source pool decreases (lower $\mathrm{E}\{p_\mathrm{n}\}$) because mutualisms introduce dependencies that prevent colonization.

As the community assembles, we find that the con-

nectance of trophic interactions ($C = \sum_{i,j} e_{\boldsymbol{A}(i,j)}/\mathcal{S}_{\boldsymbol{A}}^2$) follows a decay-like trajectory to values similar to – but on average 9% greater than – the connectance of the source pool $\boldsymbol{P}$. Decaying connectance has been documented in the assembly of mangrove communities (Peitchnik), however this decay is a statistical inevitability, as a growing food web early in the assembly process undergoes both increasing species richness and - through the establishment of different trophic levels and compartments - a decline in link density relative to the small, densely connected network at the beginning of assembly. That the connectance of assembled communities is greater than the source pool is due to the fact that only species connected by trophic interactions can enter the community to begin with, increasing expected link density compared to the overall pool.

The trophic breadth of species is thought to play an important role in community assembly, and while many measure exist, trophic generality can be defined as $G_i = \sum_j e_{\boldsymbol{A}(i,j)}(L/S)^{-1}$, such that the number of trophic interactions for a consumer is scaled by the average number of trophic interactions per species in the community L/S. A species is classified as a generalist if the number of its trophic interactions is greater than the average number of links per species, or $G_i > 1$, and a specialist if $G_i < 1$, where a community can be described by its proportion of specialists $\rho_s(G)$. Piechnik et al. scale trophic breadth to that of an average steady state value of $L^*/S^* = 0.2$, and we refer to this measure of generality as G_i^*. However, if we assume that L/S is measured with respect to current community composition, $S = \mathcal{S}_{\boldsymbol{A}}$ and $L = \sum_{i,j} e_{\boldsymbol{A}(i,j)}$. To address differences in scaling generalism against the average number of trophic links per species, we employ three different measures of L/S to calculate G_i and determine changes in the proportion of specialists in food webs over the course of assembly: 1) G_i^{all}, where L accounts for all links in the food web and S accounts for all species relative to each time interval in the assembly process (circles; Fig. 2b); 2) G_i^{hetero}, where we consider only the links and species richness of heterotrophs, excluding autotrophs (points; Fig. 2b); 3) G_i^*, where L and S are measured with respect to the communities at steady state (closest to Piechnik et al., 2008; diamonds; Fig. 2b).

Whether trophic breadth is scaled to the current state of L/S or the steady state value of L^*/S^* has a large effect on the estimated proportion of generalists in the community, particularly when the size of the system is small. We observe that for G_i^{all}, the system is initially assembled by specialist species, but where the proportion of specialists relative to generalists declines to ca. 60% by the time the community reaches steady state. If only

the trophic links between heterotrophs are considered as in G_i^{hetero}, specialists still dominate early in assembly, but there is a greater range, such that some systems can be described by a mixed proportion of specialists and generalists. If generalism is measured with respect to the steady state L^*/S^* as in G_i^*, we observe that generalists
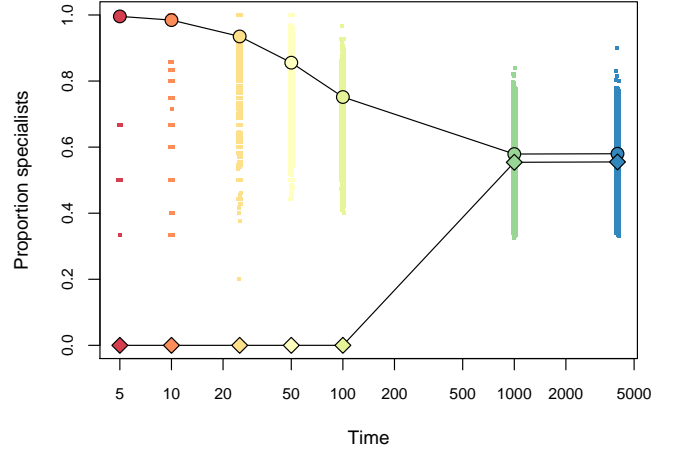

Figure 2: a) b)

dominate early in assembly.

I. DISCUSSION

The small to non-existant role of specialist species early in the assembly process if generality is measured with respect to the steady state value of L^*/S^*, as well as the increase the proportion of specialist consumers to ca. 60% based on all measures of G_i, is consistent with observations of empirical systems (Piechnick et al. 2008). However, we find the perspective of generalists dominating early assembly is very much dependent on scaling generality to the steady state L^*/S^*, and if we instead scale generality to a changing L/S over the course of assembly, it is specialists that initiate assembly, and this is particularly true if we account for both heterotrophs and autotrophs. If only heterotrophs are accounted for,

Appendices

Parameter	Definition	Value/Range
$\overrightarrow{a}$	assimilate	
$\overrightarrow{n}$	need	
$\overrightarrow{i}$	ignore	
$\overrightarrow{m}$	make	
$e \leftrightarrow i$	Asymmetric consumption	$p_{ei} = p_i(p_e/(p_e + p_n + p_i)) + p_e(p_i/(p_a + p_i + p_n))$
$e \leftrightarrow e$	Symmetric consumption	$p_{ee} = p_e(p_e/(p_i + p_n + p_e))$
$e \leftrightarrow n$	Trophic mutualism	$p_{en} = p_n(p_e/(p_e + p_n + p_i + p_m)) + p_e(p_n/(p_a + p_i + p_n))$
$n \leftrightarrow n$	Non-trophic mutualism	$p_{nn} = p_n(p_n/(p_e + p_n + p_i + p_m))$
$n \leftrightarrow i$	Commensalism	$p_{ni} = p_n(p_i/(p_e + p_n + p_i + p_m)) + p_i(p_n/(p_e + p_n + p_i))$
$m \leftrightarrow n$	Engineering	$p_{mn} = p_n(p_m/(p_e + p_n + p_i + p_m)) + p_m$
$i \leftrightarrow i$	Null	$p_{ii} = p_i(p_i/(p_e + p_n + p_i))$
$\mathcal{N}$	Number of species + objects	dyn.
$\mathcal{S}$	Number of species	dyn.
$\mathcal{O}$	Number of objects	dyn.

Table I: Table of parameters, definitions, and assigned values or ranges.